



Australian Government
Productivity Commission

Literacy and Numeracy Skills and Labour Market Outcomes in Australia

Productivity Commission
Staff Working Paper

May 2014

Anthony Shomos
Matthew Forbes

The views expressed in
this paper are those of the
staff involved and do not
necessarily reflect the views of the
Productivity Commission.

© Commonwealth of Australia 2014

ISBN 978-1-74037-487-3

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, the work may be reproduced in whole or in part for study or training purposes, subject to the inclusion of an acknowledgment of the source. Reproduction for commercial use or sale requires prior written permission from the Productivity Commission. Requests and inquiries concerning reproduction and rights should be addressed to Media and Publications (see below).

This publication is available from the Productivity Commission website at www.pc.gov.au. If you require part or all of this publication in a different format, please contact Media and Publications.

Publications enquiries:

Media and Publications
Productivity Commission
Locked Bag 2 Collins Street East
Melbourne VIC 8003

Tel: (03) 9653 2244
Fax: (03) 9653 2303
Email: maps@pc.gov.au

General enquiries:

Tel: (03) 9653 2100 or (02) 6240 3200

An appropriate citation for this paper is:

Shomos, A. and Forbes, M. 2014 *Literacy and Numeracy Skills and Labour Market Outcomes in Australia*, Productivity Commission Staff Working Paper, Canberra.

JEL code: J24.

The Productivity Commission

The Productivity Commission is the Australian Government's independent research and advisory body on a range of economic, social and environmental issues affecting the welfare of Australians. Its role, expressed most simply, is to help governments make better policies, in the long term interest of the Australian community.

The Commission's independence is underpinned by an Act of Parliament. Its processes and outputs are open to public scrutiny and are driven by concern for the wellbeing of the community as a whole.

Further information on the Productivity Commission can be obtained from the Commission's website (www.pc.gov.au) or by contacting Media and Publications on (03) 9653 2244 or email: maps@pc.gov.au

Contents

Acknowledgments	iv
Abbreviations	v
Key Points	vi
1 Introduction	1
1.1 Human capital, literacy and numeracy — why the policy interest?	2
1.2 Framework linking human capital and labour market outcomes	4
1.3 Previous research	9
1.4 Outline for the rest of the paper	10
2 A profile of literacy and numeracy skills in Australia	13
2.1 How are literacy and numeracy skills defined and measured?	13
2.2 A profile of literacy and numeracy skills in Australia	14
2.3 Literacy and numeracy skills by labour market outcomes	26
3 Modelling and results	31
3.1 Model and variables	31
3.2 Results	36
3.3 Conclusion	44
A Literacy, numeracy and problem solving measures in PIAAC	47
A.1 Proficiency measures in PIAAC	48
A.2 Measuring latent proficiencies	54
A.3 Multiple imputation and plausible values	59
B Descriptive statistics	63
B.1 Construction of variables used in the modelling	63
B.2 Descriptive statistics	65
C Estimation results	67
C.1 Labour force status models	67
C.2 Wages models	70
References	73

Acknowledgments

There were a number of Productivity Commission staff that provided valuable input to this staff working paper. The authors would particularly like to thank Lou Will and Patrick Jomini for their guidance and feedback on drafts. Jared Greenville, Jenny Gordon, Lisa Gropp, Mike Woods and former Commission staff member Noel Gaston also provided helpful comments.

The authors also wish to thank the external referees Chris Ryan (Melbourne Institute of Applied Economic and Social Research, University of Melbourne) and Garry Barrett (University of Sydney) for their helpful feedback.

Other helpful input was also provided by Bruce Caldwell and Myles Burleigh at the Australian Bureau of Statistics (ABS).

The paper used data from an ABS confidentialised unit record file from the Programme for International Assessment of Adult Competencies survey.

The findings and views reported in this paper are those of the authors and do not necessarily reflect the views of the Productivity Commission or of the external organisations and people who provided assistance.

Abbreviations

ABS	Australian Bureau of Statistics
ALLS	Adult Literacy and Lifeskills Survey
ESB	English Speaking Background
IALS	International Adult Literacy Survey
IRF	Item Response Function
IRT	Item Response Theory
LFS	Labour Force Status
ME	Marginal effect
NAPLAN	National Assessment Program — Literacy and Numeracy
NESB	Non-English Speaking Background
OECD	Organisation for Economic Co-operation and Development
PIAAC	Programme for the International Assessment of Adult Competencies
PISA	Programme for International Student Assessment
PSU	Primary Sampling Units
SAL	Survey of Aspects of Literacy
SE	Standard error

Key points

- Adult literacy and numeracy skills contribute to wellbeing in many ways. At an individual level, they are central to social and economic participation.
 - Literacy and numeracy skills are a core part of a person's human capital.
 - They also support the development of other forms of human capital, including knowledge, other skills and health.
- Some Australians have low (level 1 or below) literacy and numeracy skills. In 2011–12:
 - 14 per cent of Australians could, at best, read only relatively short texts from which they were able to locate only a single piece of information.
 - 22 per cent could only carry out one-step or simple processes such as counting where the mathematical content is explicit with little or no text or distractors.
- At the other end of the skill distribution, 16 per cent of Australians had high (level 4/5) literacy skills and 12 per cent had high numeracy skills in 2011–12.
 - People with high literacy skills can make complex inferences and evaluate subtle truth claims or arguments in lengthy or multiple texts.
 - People with high numeracy skills can understand a broad range of mathematical information that may be complex, abstract or embedded in unfamiliar contexts.
- Most Australians have skills somewhere between these levels. Groups with relatively low literacy and numeracy skills include: people with low levels of education; older persons; people not working; and immigrants with a non-English speaking background.
- Compared with other countries in the OECD, Australia performs above average on literacy but average in numeracy.
- Higher literacy and numeracy skills are associated with better labour market outcomes (employment and wages). Econometric modelling shows that:
 - an increase in literacy and numeracy by one skill level is associated with an increased likelihood of employment of 2.4 and 4.3 percentage points for men and women, respectively
 - an increase in literacy and numeracy skills is associated with a similar increase in the probability of employment, whether a person had a degree, diploma/certificate or Year 12 education
 - an increase in literacy and numeracy by one skill level is associated with about a 10 per cent increase in wages for both men and women. This positive association is equivalent to that of increasing educational attainment from Year 11 to Year 12 or to a diploma/certificate
 - up to 40 per cent of the association between education and employment is attributable to literacy and numeracy skills. These results are consistent with education providing many other attributes of human capital that are valued in the workplace
 - more than half of the 'penalty' that affects the wages of people with a non-English speaking background is explained by their lower literacy and numeracy skills.

1 Introduction

Literacy and numeracy skills form part of a person's 'human capital', and are important for economic and social participation. Research for Australia has found that having better literacy and numeracy skills increases the likelihood of positive labour market outcomes (for example, Chesters, Ryan and Sinning 2013; Barrett 2012).

This paper profiles the literacy and numeracy skills of Australia's adult population and assesses how important they are for two labour market outcomes — employment and wages. Results confirm findings from previous research. Specifically, they show that:

- many people have relatively low literacy and numeracy skills and the types of literacy and numeracy tasks they can do are limited in comparison with people who have higher skills. For example:
 - 14 per cent of Australians aged 15–74 (2.4 million people) have low literacy (level 1 or below) meaning they can, at best, read only relatively short texts from which they can locate only a single piece of information (detailed descriptions of tasks for each skill level are in table A.1).
 - 16 per cent of the population have high literacy (level 4/5), meaning they can make complex inferences and evaluate subtle truth claims or arguments in lengthy or multiple texts.
- there is a high correlation between a person's literacy and numeracy skills
- literacy and numeracy skills vary across different groups. On average:
 - people with a non-English speaking background have lower (English) literacy and numeracy skills than other people
 - older persons (55–74) have lower literacy and numeracy than younger persons
 - more highly educated people have higher literacy and numeracy
- there is a strong positive association between literacy and numeracy skills and labour market outcomes.

Section 1.1 of this chapter describes the recent policy focus on human capital and on improving literacy and numeracy. Section 1.2 develops a framework for

understanding how literacy and numeracy skills are developed and the relationship between those skills and labour market outcomes. Section 1.3 summarises findings from previous research that has examined the association between human capital (including literacy and numeracy) and labour market outcomes, and outlines the rest of the paper.

1.1 Human capital, literacy and numeracy — why the policy interest?

Human capital can be defined as ‘the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being’ (OECD 2001, p. 18).

People with more human capital tend to enjoy better health, improved life satisfaction and higher levels of social engagement (McLachlan, Gilfillan and Gordon, 2013).

People with more human capital are also more productive. Investment in education and training¹ increases a person’s productivity and his or her gross returns² from working, as measured by wages (Becker 1993). As people acquire more human capital, they are more likely to enter the workforce and earn more, all else equal.

Literacy and numeracy skills are becoming increasingly important for productivity, as they provide the foundation to develop other skills:

The basic skills acquired in early childhood and school years, particularly literacy and numeracy, are the necessary foundation for developing higher order skills that contribute to a more productive workforce. (Treasury 2010)

... the demands of the ‘information age’ increasingly require higher level skills that are best acquired through formal education and training. Such skills are of two kinds: specific and generic. Both are important, but the innovation and adaptation that underpin productivity growth are placing increasing demands on the more general analytical, discovery and communication skills. These are grounded in the literacy and numeracy acquired progressively at school and developed through higher education. (Banks 2012, p. 11)

Reflecting the importance of these skills for adults to fulfil their potential, there are several government programs that focus on improving and monitoring literacy and numeracy of various demographic groups (box 1.1).

¹ This investment can be made by the person, an employer or the government.

² Net returns from education and training are less than gross returns, as they take account of the cost of training and any lower earnings during the investment period.

Box 1.1 Government programs targeting literacy and numeracy

The National Foundation Skills Strategy for Adults

The Strategy has a target to increase the foundation skills (defined in the Strategy as language, literacy and numeracy and employability skills) of persons aged 15–64. The initial emphasis of the Strategy is on people with lower level skills, as the greatest economic impact on labour productivity can be gained from improving skills at lower levels (SCOTese 2012).

The Skills for Education and Employment Program

This program ‘provides language, literacy and numeracy training to eligible job seekers, with the expectation that such improvements will enable them to participate more effectively in training or in the labour force’ (Department of Industry 2013b).

Closing the Gap — Expansion of intensive literacy and numeracy programs

This initiative builds upon existing teaching and learning practices of literacy and numeracy for Aboriginal and Torres Strait Islander students (DEEWR 2013).

Language, Literacy and Numeracy (LLN) Practitioner Scholarships Program

This program seeks to address skill shortages in the adult LLN field in Australia by providing financial incentives to increase the number of qualified LLN practitioners, particularly in regional areas (Department of Industry 2013a).

The National Assessment Program — Literacy and Numeracy (NAPLAN)

This is an annual assessment of reading, writing, language and numeracy for students in years 3, 5, 7 and 9. NAPLAN started in 2008 and tests skills that are essential for children to progress through school. NAPLAN results are used to determine student and school performance (ACARA 2013).

Given the recent focus on improving literacy and numeracy outcomes, it is important to understand:

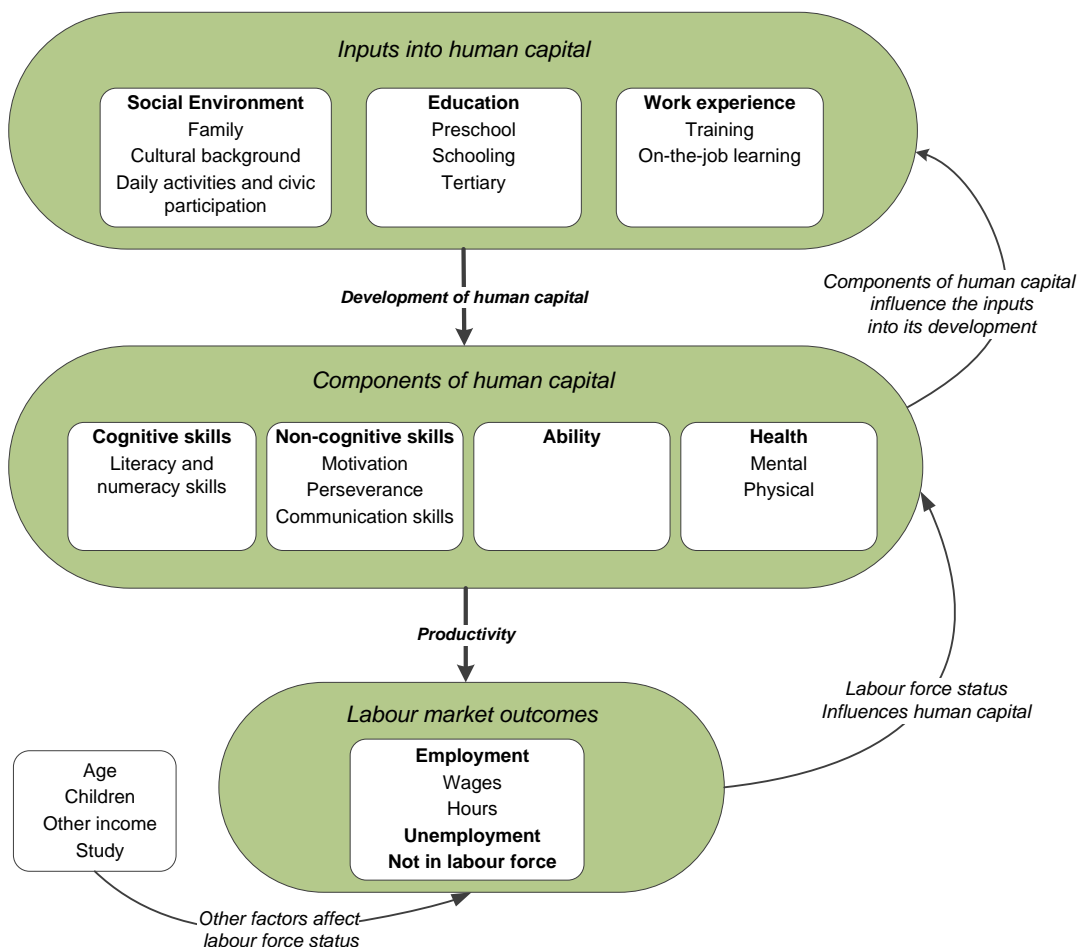
- the profile of literacy and numeracy skills in the population, and how they relate to people’s educational and other characteristics
- how the skills of the population (and of various demographic groups) change over time
- the relationships between literacy and numeracy skills and labour market outcomes, as these outcomes represent a large part of the private and social expected returns to the investment in improved literacy and numeracy skills.

This paper seeks to address these issues. In particular, the aim of the paper is to examine how changes in literacy and numeracy skills are associated with changes in employment and wages.

1.2 Framework linking human capital and labour market outcomes

This section develops a framework of how human capital is created, and how it influences labour market outcomes (figure 1.1). Other relevant relationships are also drawn out as required.

Figure 1.1 **Analytical framework for human capital and labour market outcomes**



As a stylised model, the figure is necessarily a simplification of the concepts and processes it represents. The model consists of three parts: the way in which human capital is developed, the components of human capital, and how they affect labour market outcomes along with other influences on these outcomes.

What is human capital?

Human capital encompasses many different attributes, which are represented in the middle of figure 1.1.

Cognitive skills

These include capacities used to process information, including memory and perception. Literacy and numeracy skills are part of a person's cognitive skills, and are sometimes referred to as 'foundation' skills, as they provide the basis to develop other, higher order skills. (Definitions of literacy and numeracy used in this paper are in chapter 2.)

Non-cognitive skills

Heckman and Kautz (2013) describe these as 'character' skills, and refer to various attributes, including perseverance, self-control, trust, self-esteem, resilience in adversity and the ability to engage productively in society. They argue that these skills are malleable until later ages and can be shaped through schooling and in other settings, but also report on a large body of evidence that shows stability in character skills.

In a separate study, Cobb-Clark and Schurer (2011) found that 'personality' as measured by the 'big five personality traits' — extraversion, agreeableness, conscientiousness, emotional stability and openness to experience — is very stable over time.

Ability

The concept of ability is closely related to the notion of skill, but in contrast to skill, ability is a latent or unobserved capability. The potential for a latent ability to affect academic success and labour market outcomes is important when estimating their relationships with cognitive skills. For instance, it can bias upwards the observed returns to schooling because people with higher cognitive ability may find it easier to undertake education (Leigh 2008).

Health

This includes the physical, emotional and mental health of a person. The mental health of a person can also affect his or her capacity to learn and to apply skills and abilities.

Development of human capital

Informal learning and the skills and knowledge acquired in formal education occur throughout a person's life. The Organisation for Economic Co-operation and Development (OECD, 2001) considers that human capital is developed by:

- learning within the family and in early childcare settings
- informal learning 'on-the-job' and in daily living and civic participation
- formal education and training (including early childhood, school and tertiary education)
- workplace training and other learning at work, through activities such as research and innovation or participation in professional networks.

Many 'inputs' (depicted at the top of figure 1.1) enter into the development of human capital.

Inputs into human capital

The *social environment* (family and cultural background, and daily activities) can influence learning and the development of human capital.

The learning and knowledge acquired through *education* develops human capital. Different types of education are likely to be more or less important for developing certain types of skills and knowledge.³ For example, foundation skills such as literacy and numeracy are more easily developed through early schooling and are more difficult for adults to learn (AWPA 2013). Some occupations require qualifications that are obtained through formal education.

Work experience develops on-the-job skills, and may help maintain a person's skills. For example, Desjardins and Warnke (2012, p. 5) stated that while 'skills can continue to increase as a function of work experience, they may also depreciate due to a lack of use'. Some skills might be firm-specific, whereas general skills, such as literacy and numeracy skills, are more transferable to other jobs.

³ This paper does not seek to determine the relative importance of different inputs into developing human capital.

Different human capital components are interdependent

Different components of human capital influence and reinforce each other. For example, people with high literacy understand written materials easily, which allows them to develop knowledge and other higher order skills.

There is an abundant literature about the links between education and healthy lifestyles. The OECD (2010a, 2012a) argue that improving educational attainment can promote healthier lifestyles. Cutler and Lleras-Muney (2006) find that better educated people are less likely to smoke, drink alcohol excessively, be obese or use illegal drugs. Although they cannot conclude why education influences health, the authors think it is likely, in part, because education improves cognitive skills. These skills include better health literacy and knowledge, or decision making skills. For example, the ‘more educated do appear to be better informed, and appear to make use of new health related information first’ (p. 16). The positive correlation between human capital and health status could be due to a third factor, such as a person’s rate of time preference. People who invest in human capital are likely to have lower discount rates because the benefits of education are somewhat delayed (compared with working). Such people are also more likely to adopt positive health behaviours with long-term benefits. Cutler and Lleras-Muney (2006) cite research that shows more educated persons have lower discount rates, although the relationship is weak.

A person’s human capital can also influence the inputs that develop human capital

The stock of human capital can influence whether a person chooses to develop it further, and the gains they get from further development.

For example, because people with better literacy find it easier to understand material presented during formal education, they may gain more benefits from additional education, which makes them more likely to undertake more of it.

Similarly, people with greater motivation and perseverance may choose to undertake more education. As measures of these abilities are rarely included in datasets it is difficult to identify their influence on education (and labour market outcomes).⁴

⁴ With longitudinal data, fixed and variable effects models can be used to allow for unobserved individual heterogeneity (such as ability) that is either innate or varies with life experience.

Human capital interacts with other factors to determine labour market outcomes

The main labour market outcomes of interest in this paper are employment⁵ and wages. These outcomes are determined by the interaction of human capital with other factors, as represented in the lower part of figure 1.1.

How human capital influences labour market outcomes

People with higher ability, skills, and knowledge, can produce more highly valued output than people with lower skills, knowledge and ability can. People who are more productive in this sense are more likely to achieve a higher wage,⁶ which increases their likelihood of entering the workforce and gaining employment.

Similarly, health influences labour force participation because it determines the amount of time a person can spend producing output (Grossman 2000). Poor health can also reduce productivity at work from increased absenteeism or from not being fully effective at work (Medibank 2011).

To summarise, human capital improves labour productivity, which in turn makes it more likely that a person will gain employment and earn higher wages.

Other factors influence labour market outcomes

A person's labour force status is also determined by factors other than human capital. These can include people's age, whether they have children, or are studying.⁷ For example, older persons may reduce their hours of work or retire. Birch (2005) found that the presence of children is likely to reduce a woman's labour supply because children increase the demand for non-work activities. Compared with having younger children, as children get older they require less parenting time and place additional demands on household budgets which can increase the chances of women participating in the labour market. Birch (2005)

⁵ Labour force states other than employment (unemployed and not in the labour force) are also considered in chapters 2 and 3.

⁶ According to human capital theory (Becker 1993), people who invest in training earn a higher wage after that training is completed. However, their earnings while training are often lower as they allocate part of their time to studying.

⁷ Some of these factors could also impact on human capital. As some people get older, their skills can depreciate because, for example, those skills are not used (see Desjardins and Warnke 2012 for a review of the literature on ageing and skills).

shows that the majority of Australian research indicates that the presence of young children (less than 6 years old) substantially reduces the labour supply of women.

The relationship between human capital and labour market outcomes might not be one-way

Although human capital influences labour market outcomes, those outcomes can in turn influence human capital. Labour force status has been linked to health outcomes and vice versa (Cai and Kalb 2006; Laplagne, Glover and Shomos 2007; Cai 2010). Employment could also help maintain skills required in the workplace (including literacy and numeracy skills).

1.3 Previous research

Most of the literature on the relationships between human capital and labour market outcomes has used measures of education and work experience to proxy for human capital, and has not included measures of specific skills. That research confirms a positive effect of education (and experience) on labour market outcomes for Australia (for example, Laplagne, Glover and Shomos 2007, for labour force participation and Forbes, Barker and Turner 2010 and Leigh 2008, for wages).

As noted above, education is only one input into the development of literacy and numeracy skills and other components of human capital. Importantly, educational attainment is not a direct measure of human capital. The OECD (2001, p. 10) argues that education ‘is a poor indicator of the stock of human capital. People with the same nominal level and type of education can differ markedly in their command of various skills’. Therefore, research that uses education as a proxy for human capital might not produce accurate estimates of how human capital (including literacy and numeracy, ability and health) influences labour market outcomes.

In Australia, the Australian Bureau of Statistics has released three surveys that include direct assessments⁸ of the literacy and numeracy skills of Australian adults (aged 15–74):

- Survey of Aspects of Literacy (SAL) (1996)
- Adult Literacy and Lifeskills Survey (ALLS) (2006)

⁸ Although other large surveys for Australia, such as the Household, Income and Labour Dynamics of Australia (HILDA) survey, contain information on skills, they typically use subjective measures (for example, self-assessment). The 12th wave of HILDA includes tests of cognitive skills and abilities (Wooden 2012).

-
- Programme for the International Assessment of Adult Competencies (PIAAC) (2011–12).

Research based on the first two surveys confirms a positive association between literacy and numeracy skills and labour market outcomes. Chiswick, Lee and Miller (2003) examined the effect of literacy and numeracy on labour force participation using the SAL data. The effects of literacy and numeracy skills on labour market outcomes were found to be quite large compared with the effects of education. For example, the estimated impact on male unemployment from changing literacy and numeracy skills from the lowest to the highest levels was estimated to be equivalent to the effect of 17 years of education.

Using the 2006 ALLS data, Barrett (2012) found that men's literacy was positively related with wages. A one standard deviation increase in literacy was associated with a 17 per cent increase in wages. Shomos (2010) also found a positive association between literacy and numeracy skills and labour market outcomes (wages and labour force participation). An increase in literacy and numeracy skills from level 1 to level 3 (which is about a two standard deviation change in skills) was associated with an increase in hourly wages of 25 and 30 per cent for women and men, respectively.

Using Australian data for both 1996 and 2006, Chesters, Ryan and Sinning (2013) found that returns from literacy were highest for workers with high levels of education over that period. That is, the positive association between literacy and income was larger for more educated people. They also found that, over the ten year period, returns to literacy had increased for young workers who had low to medium levels of education.

Another important finding from these papers is that, even after controlling for literacy and numeracy skills, the association between education and labour market outcomes is still positive and significant, consistent with education contributing to the formation of many other skills deemed to be valuable in the workforce.

1.4 Outline for the rest of the paper

This paper uses 2011–12 data from PIAAC to estimate the association between literacy and numeracy skills and labour market outcomes (employment and wages). The focus of the analysis is for Australia. Other studies that the authors are aware of that have used the PIAAC data are by the OECD (2013a) and Hanushek et al. (2013). Results in this paper for Australia are not directly comparable with those produced by the OECD (2013a) because the OECD used only literacy in its analysis

whereas this paper uses a measure that combines both literacy and numeracy.⁹ However, the findings here are consistent with the other results (a comparison is provided in chapter 3). Hanushek et al. (2013) pooled results for 22 countries, but did not include Australia in their analysis.

The paper builds on previous research that used data from 2006. Each study has been conducted at a different point in the economic cycle, making it possible that the relationship between skills and labour market outcomes differ according to the environment that conditions labour market outcomes.

Chapter 2 contains a profile of literacy and numeracy in Australia for 2011–12. Descriptive analysis is used to examine how literacy and numeracy skills vary across personal characteristics — including age, educational attainment and English speaking background. Identifying the characteristics associated with lower skills can inform analyses of where the largest potential gains might be from improving skills.

The cross-tabulations in chapter 2 illustrate how literacy and numeracy skills vary by labour market outcomes. This analysis is not sufficient to determine how important literacy and numeracy skills are for wages and employment. For example, factors such as education and age are associated with literacy and numeracy skills, and are also likely to influence labour market outcomes. Furthermore, the causation between many variables is not unidirectional (figure 1.1), meaning that an association between two variables does not explain the effect of one variable on another.

To address some of these issues, chapter 3 uses multivariate analysis to isolate the association between literacy and numeracy and education from that of other factors that influence labour market outcomes. This allows for more robust estimates of the relationship between literacy and numeracy skills, and labour market outcomes, than can be drawn from cross-tabulations. Furthermore, measuring the association between educational attainment and labour market outcomes when literacy and numeracy are taken into account gives an indication of the importance of other skills that are associated with education.

The modelling does not account for potential reverse causality — for example, the effect of labour force status on literacy and numeracy skills. The model also cannot account for some factors that are difficult to observe, and which are likely to be linked to both labour market outcomes and literacy and numeracy skills (for

⁹ The measures were combined because they were highly collinear (chapter 2 and appendix B).

example, ability).¹⁰ Therefore, results are interpreted as associations and evidence that supports the effect of one factor on another, as hypothesised in the conceptual model (figure 1.1), bearing in mind the causality issues identified.

Chapter 3 concludes with a discussion of the results. It is found that higher literacy and numeracy skills are associated with higher employment and wages. There is also a positive association between education and labour market outcomes. Education is important in developing literacy and numeracy skills as well as other components of human capital valued in the workplace.

¹⁰ Longitudinal data that observes outcomes for persons over time could be used for this analysis but PIAAC is a cross-sectional data source.

2 A profile of literacy and numeracy skills in Australia

The measures of literacy and numeracy skills used in the analysis are outlined in section 2.1. Australians' literacy and numeracy skills vary across demographic groups (section 2.2) and labour market outcomes (section 2.3).

2.1 How are literacy and numeracy skills defined and measured?

The Australian Bureau of Statistics (ABS) conducted the Programme for International Assessment of Adult Competencies (PIAAC) survey during 2011–12 on behalf of the Organisation for Economic Co-operation and Development (OECD). The survey has been conducted across 23 countries and the Russian Federation. Respondents were given various tasks to assess their skills in three domains: literacy; numeracy; and problem solving in a technological environment. The focus in this paper is on the literacy and numeracy skill domains, where:

- literacy is defined as 'understanding, evaluating, using and engaging with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential' (OECD 2012b, p. 20), and
- numeracy is defined as 'the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life' (OECD 2012b, p. 34).

For each domain, each person's skill was estimated with a test score ranging from 0–500. Five skill levels were defined within this range (table 2.1). Appendix A contains more information on the derivation of test scores and associated skill levels.

Few people attain skill level 5; this can result in large standard errors for estimates of this part of the population. As a result, the ABS typically combines skill levels 4 and 5. This approach is also adopted in this paper. Level one and below level one

are also combined in presenting results, although the quantitative analysis itself is conducted with the test scores.¹

Table 2.1 Concordance of test scores to skill levels for literacy and numeracy

<i>Test score</i>	<i>Skill level</i>
0 to < 176	Below level 1
176 to < 226	1
226 to < 276	2
276 to < 326	3
326 to < 376	4
376 to 500	5

Source: Appendix A.

2.2 A profile of literacy and numeracy skills in Australia

In 2011–12, 14 per cent of the population had relatively low literacy skills (at or below level 1) (table 2.2). These people could, at best, read only relatively short texts from which they were able to locate only a single piece of information. In contrast, 16 per cent of the population had high (level 4/5) literacy skills. These people can make complex inferences and evaluate subtle truth claims or arguments in lengthy or multiple texts. Skills are measured on a continuum, and the majority of the population have literacy skills somewhere in between these levels.

In numeracy, 22 per cent of the population had skills at or below level 1. These people could carry out one-step or simple processes such as counting where the mathematical content is explicit with little or no text or distractors. At the other end of the distribution, 12 per cent of the population had relatively high (level 4/5) numeracy skills. People with high numeracy skills can understand a broad range of mathematical information that may be complex, abstract or embedded in unfamiliar contexts.

Descriptions of the types of tasks that correspond to each skill level are in tables A.1 and A.2 for literacy and numeracy, respectively.

¹ The quantitative analysis was also based on 10 ‘plausible values’ for each person’s literacy and numeracy skills. This required ‘Rubin’s rules’ to compute test scores for literacy and numeracy taking into account each plausible value (appendix A describes this process).

Table 2.2 Literacy and numeracy skill levels

Share of population aged 15–74, per cent^a

	<i>Level 1 or below</i>	<i>Level 2</i>	<i>Level 3</i>	<i>Level 4/5</i>
Literacy	14.1	30.1	37.9	15.6
Numeracy	21.7	32.5	31.3	12.3

^a Totals do not add to 100 per cent because some observations were labelled as missing.

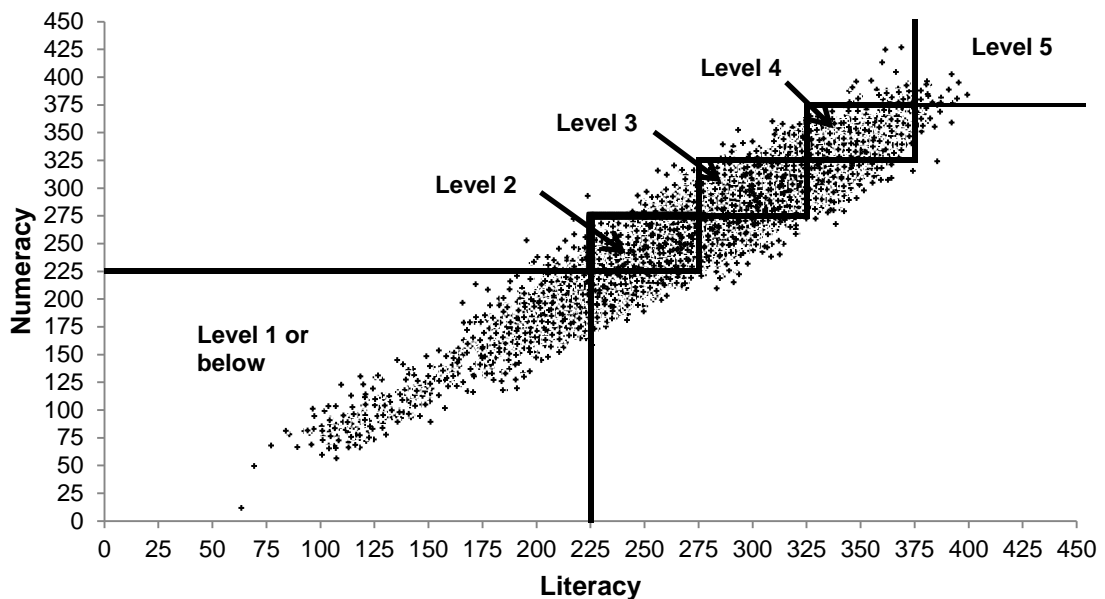
Source: ABS (2013a).

Correlations

There is a close relationship between proficiency in literacy and numeracy. The correlation between literacy and numeracy test scores for persons in the entire sample of the PIAAC data is 0.90 (figure 2.1). The high correlation between the measures meant that the analysis in chapter 3 is based on an average of both test scores.

Figure 2.1 Correlation between literacy and numeracy test scores

15–74 year olds



Source: Authors' estimates based on PIAAC data.

When the data are analysed in terms of skill level, more than 60 per cent of the population demonstrated the same measured skill levels for literacy and numeracy (table 2.3). Of the remaining population, most achieved relatively higher literacy scores — 29 per cent of people had a literacy skill level which was one level higher

than their numeracy skill level. In contrast, fewer than 10 per cent of people were assessed with a lower literacy skill level than their numeracy skill level.

Table 2.3 Correlation between literacy and numeracy, by skill level^a
Per cent of 15–74 year olds

<i>Literacy</i>	<i>Numeracy</i>				Total
	Level 1 or below	Level 2	Level 3	Level 4/5	
Level 1 or below	13	1	0	0	14
Level 2	9	18	4	0	31
Level 3	1	13	22	4	39
Level 4/5	0	0	7	9	16
Total	22	32	33	13	100

^a Totals may not add due to rounding.

Source: Authors' estimates based on PIAAC data.

How does Australia compare internationally?

The PIAAC survey is designed to allow comparisons across all countries that participated in the survey. Participating countries were responsible for translating both the instruments used to assess literacy and numeracy skills and the background questionnaire (box 2.1).

As the survey is designed to gauge a person's ability to participate in the society in which they live, the tests were conducted in the official language of the country of residence. In Australia, for persons with a non-English speaking background PIAAC does not measure proficiency in first language.

Box 2.1 **Comparability of PIAAC data across countries**

Australia was one of 24 countries that participated in the Programme for International Assessment of Adult Competencies (PIAAC).

Who was surveyed?

The target population for the survey was the non-institutionalised population, aged 16–65 years, residing in the country at the time of data collection, irrespective of nationality, citizenship or language status. In Belgium, only the Flemish region (Flanders) participated in the survey. In the United Kingdom, only the autonomous administrative regions of England and Northern Ireland participated in the study.

Some countries oversampled, while others sampled individuals outside the target population. For example, Australia included 15 year olds and persons aged 66–74.

How were the instruments for assessment translated across countries?

The translation and adaptation of instruments was a core component of how the survey was constructed.

Participating countries were responsible for translating the assessment instruments and the background questionnaire. Any national adaptations of either the instruments or the questionnaire was subject to strict guidelines, and to review and approval by the international consortium. The recommended translation procedure was for a double translation from the English source version by two independent translators, followed by reconciliation by a third translator (OECD 2013c, p. 55).

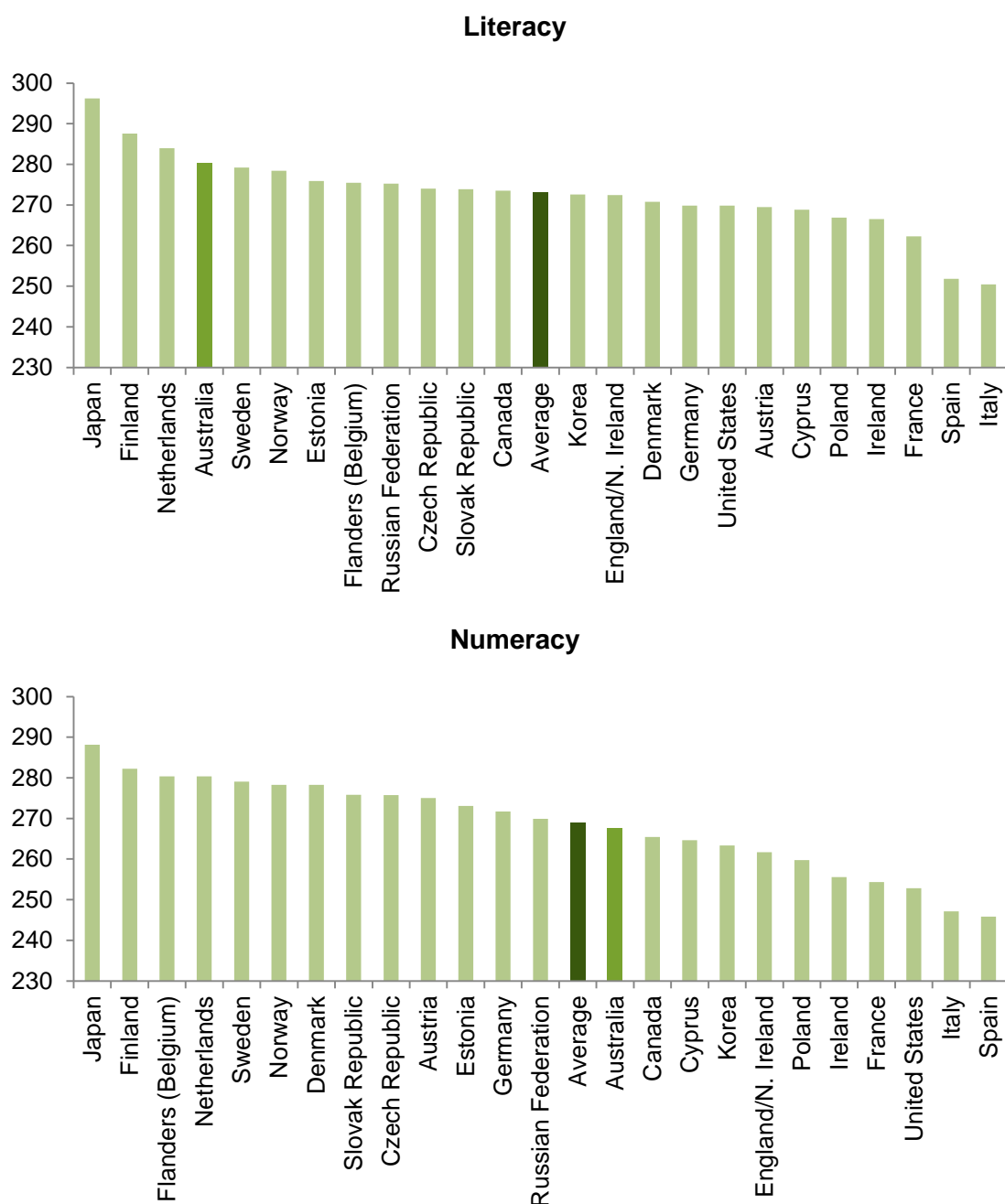
Literacy-related non-response

Some respondents were unable to undertake the assessment for literacy-related reasons, such as being unable to speak or read the test language(s), having difficulty reading or writing, or having a learning or mental disability. Some of these respondents completed the background questionnaire, or its key parts, presumably with the assistance of an interviewer or family member. The available background information regarding these respondents was used to impute proficiency scores in literacy and numeracy.

Source: OECD (2013c).

Most countries' ranking in literacy and numeracy are similar, although there are some notable exceptions, with Australia being one of them. Australia performed significantly above average in literacy, but its performance in numeracy was similar to the average (figure 2.2). It is not clear why Australia performs better in literacy, whereas most other countries perform similarly in both domains.

Figure 2.2 Literacy and numeracy test scores, by country, 2011–12^{a,b}
16–64 year olds



^a Data for the Russian Federation are preliminary and subject to change. ^b The average does not include results for the Russian Federation.

Source: OECD (2013a, 2013b).

Reasons for performance differences across countries are complex and ‘likely to be affected by such factors as the historical patterns of participation in education, support for adult learning, and patterns of immigration’ (OECD 2013a, p. 74). Some of these factors are considered in more detail for Australia in the next section.

Differences in literacy and numeracy skills by demographic groups

There are many obstacles to comparing survey data over time in detail (box 2.2). As a result, this paper does not examine changes in literacy and numeracy skills by different demographic groups over time. Rather, the analysis focusses on differences in literacy and numeracy skills across demographic groups in 2011–12.

Box 2.2 Have literacy and numeracy skills changed since 1996?

The Programme for the International Assessment of Adult Competencies (PIAAC), Adult Literacy and Lifeskills Survey (ALLS) and Survey of Aspects of Literacy (SAL) have similar measures of literacy and numeracy skills. However there are some differences that must be accounted for when comparing the data over time. The Australian Bureau of Statistics (ABS, 2013b) have made the following adjustments in publishing comparable estimates:

- The literacy and numeracy scores previously published for ALLS and SAL were based on a model with a response probability value of 0.8. These scores have been re-estimated with a response probability value of 0.67 to be consistent with PIAAC.
- The ALLS and SAL had measures for two types of literacy skills — prose literacy and document literacy. These have been combined to produce a single proficiency scale comparable with literacy in PIAAC.
- The numeracy scores from the ALLS have been recalculated using a model that incorporates results of all countries that participated in ALLS which led to some minor changes in ALLS numeracy scores. SAL did not collect a comparable numeracy domain.

After these adjustments, the ABS (pers. comm., 27 February 2014) indicates that the literacy test scores show a slight decline since 2006. Similarly, the Organisation for Economic Co-operation and Development (OECD, 2013b) reports a decline in numeracy across most countries. The ABS is not confident that these measured declines are real. Analysis done by the ABS and internationally has shown that in some cases the observed trend is difficult to reconcile with other information (ABS 2013b). It could be an artefact of differences in testing processes in different years. The OECD has a small working group examining this issue.

(Continued next page)

Box 2.2 (continued)

Despite these issues, the OECD (2013b) reports a small (but statistically significant) increase in average literacy over the period 1996 to 2012 for Australia (see table below). The change in literacy scores between 2006 and 2011–12 of 13 points is about one quarter of the 50 point range for most skill levels. In contrast to literacy, average numeracy scores were reported as remaining unchanged between 2006 and 2011–12.

Average literacy and numeracy test scores over time

16–64 year olds, 1996–2012^a

	1996	2006	2011–12
Literacy	267.7 (1.0)	273.0 (0.6)	280.4 (0.9)
Numeracy	-	267.9 (0.7)	267.6 (1.0)

^a Standard errors in parentheses.

Source: OECD (2013b).

Men have higher numeracy skills than women

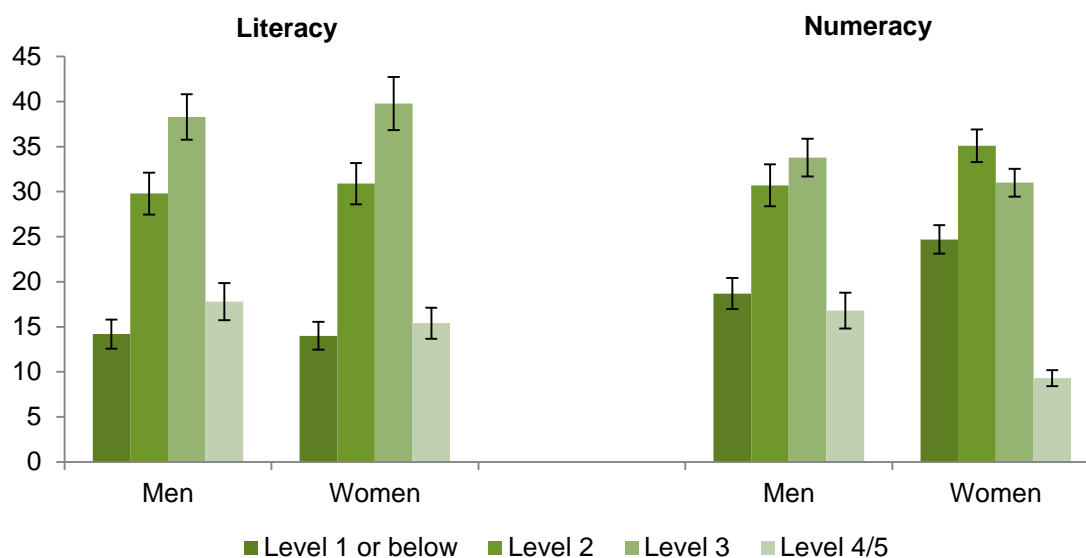
About 25 per cent of women had numeracy at skill level 1 or below in 2011–12, compared with about 17 per cent of men (figure 2.3).² The proportion of men assessed with the highest level of numeracy (level 4/5) was 17 per cent, nearly twice that of women (9 per cent).

In contrast, there were similar proportions of men and women assessed at each literacy skill level.

This pattern is the same as found in the 2006 data (Shomos 2010). The OECD (2013a) also reported that men have higher average numeracy scores than women for all surveyed countries.

² Some of the analysis in this chapter is done by comparing the discrete skill levels rather than test scores. The qualitative findings are consistent regardless of which measure is used.

Figure 2.3 Literacy and numeracy skill levels, by gender, 2011–12^a
Per cent, 15–74 year olds



^a Error bars indicate 95 per cent confidence intervals.

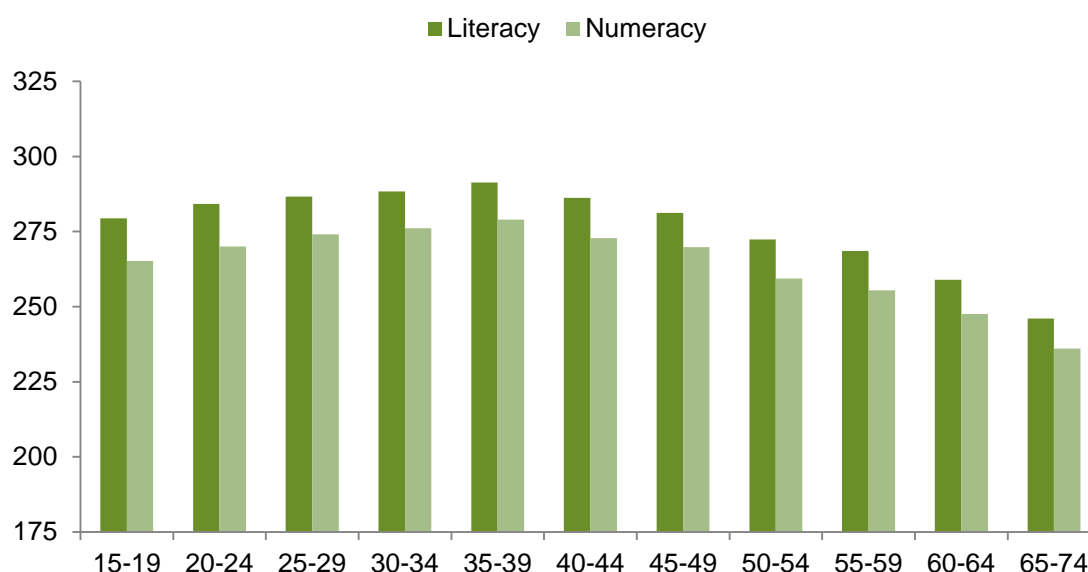
Source: ABS (2013a).

Older people have lower literacy and numeracy skills

Literacy and numeracy skills are likely to vary across age cohorts: they can be acquired, maintained or diminish at various rates during the life cycle (Statistics Canada and OECD 2005). People develop skills through formal education mostly when they are young. They also acquire skills in a multitude of settings over their lifetimes. But there has been an increase in educational attainment beyond what might be expected from life cycle factors over the past few decades.

By age cohorts, average literacy and numeracy scores increase until the age of 40 years (figure 2.4). After age 40, there is a decline in literacy and numeracy, with the old (aged 65–74 years) having the lowest literacy and numeracy.

Figure 2.4 Literacy and numeracy test scores, by age cohort, 2011–12



Source: Authors' estimates based on PIAAC data.

Others have found a similar pattern of skill variation across age cohorts. For example, Chesters, Ryan and Sinning (2013) found 'there is something of a decline in average [document literacy] scores after middle age for both men and women' (p. 16) in the 1996 and 2006 surveys.

Lower scores reported for older persons do not necessarily imply that skills deteriorate with age. For example, younger people might have undertaken more education than older people did at the same age and be more highly skilled because of that education. With only one cross-section of data it is not possible to distinguish the longitudinal effect of ageing on literacy and numeracy from the pattern of difference in literacy and numeracy by birth-cohort.

However, previous research has shown that skills are likely to deteriorate with age. Chesters, Ryan and Sinning (2013) use a 'synthetic' cohort³ analysis with data for 1996 and 2006 to examine how the skills of one birth-cohort change over time. They find the skills of each cohort older than 35–39 years decline with age. For example, the average literacy score of the cohort of men aged 40–44 years in 1996 was 21 points lower in 2006 (when they were aged 50–54 years). Green and Riddell (2007), using Canadian data, also find that skills deteriorate with age.

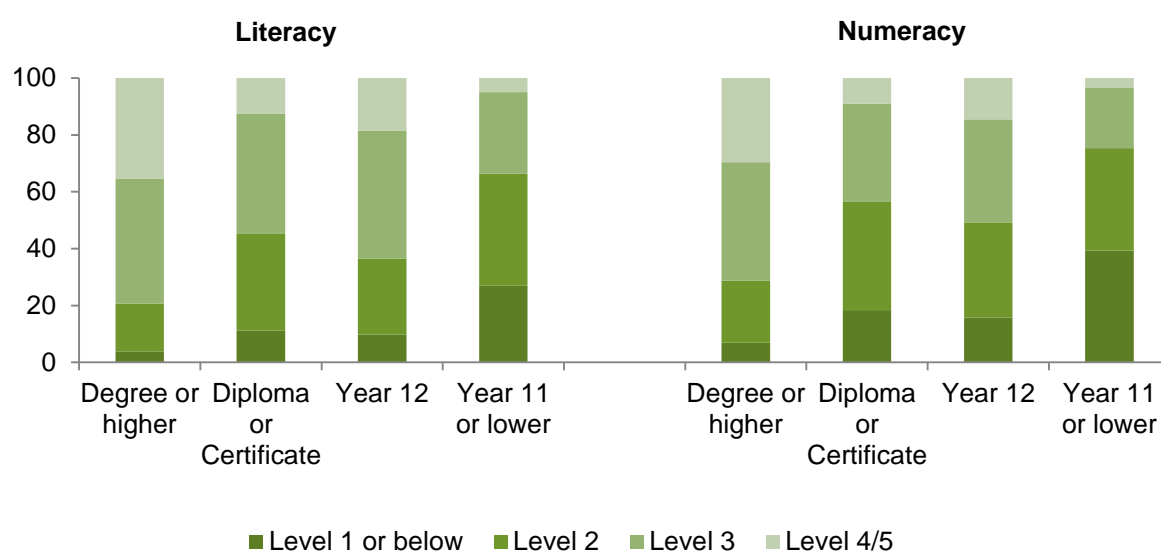
³ A cohort analysis follows the same persons across time, allowing a comparison of their test scores at different times. However, the PIAAC and predecessor surveys are for a different group of persons, meaning that a change in a person's test score cannot be estimated. A synthetic cohort assumes that a group of people at the same age in 2006 can be represented by a group that is 6 years older when measured in 2012.

There is a positive relationship between literacy and numeracy and education

Education has been shown to be important for developing and enhancing literacy and numeracy skills (Green and Riddell 2012). Furthermore, as discussed in chapter 1, having higher ability or skills could influence a person's decision to undertake more education.

The PIAAC data confirm a positive association between educational attainment and skills (figure 2.5). Differences in skills for people with the highest and lowest levels of educational attainment are large — about 4 per cent of people with at least a degree had literacy skill level 1 or below, compared with 27 per cent of people who had year 11 or lower education.

Figure 2.5 Distribution of literacy and numeracy skills by educational attainment, 2011–12
Per cent, 15–74 year olds



Source: Authors' estimates based on PIAAC data.

One might expect people with a diploma/certificate to have higher literacy and numeracy skills than people with Year 12, as the former is defined⁴ for the purposes of this study as a higher qualification. However, people with year 12 or diploma/certificate attainment display similar literacy profiles — about 10 per cent for each of these groups had literacy level 1 or below. The results could reflect selection effects (for example, choosing to do a certificate or diploma rather than

⁴ 'Diploma or certificate' is defined in this paper to include people with certificate III/IV. People with certificate I/II are defined as having year 11 or lower education (appendix B).

Year 12 because of practical skills). This cannot be controlled for, as the data do not allow for comparing the education paths each person takes. For example, a person with Year 11 and a diploma/certificate cannot be distinguished from a person with Year 12 and a diploma/certificate (appendix B).

The profile for numeracy skills by educational attainment is similar to the profile for literacy skills. However, the proportion of people with lower numeracy skills is higher across each level of educational attainment, consistent with the lower levels of numeracy in the population (figure 2.3).

Shomos (2010) (using 2006 data) and Chesters, Ryan and Sinning (2013) (with 1996 and 2006 data) also found a positive association between literacy and educational attainment.

Chesters, Ryan and Sinning (2013) also reported that both men and women in 1996 had higher, on average, document literacy skills than their counterparts with an equivalent level of educational attainment in 2006. Average literacy scores were unchanged, however, because more people obtained higher levels of educational attainment in 2006.

A similar analysis of changes in literacy skills by educational attainment since 2006 cannot be made because of the issues regarding the differences between the different surveys mentioned above. However, educational attainment increased between 2006 and 2011–12⁵ and, at an aggregate level, any improvement in literacy skills is likely to be small and numeracy skills remained relatively unchanged (see discussion regarding comparability of literacy and numeracy over time in box 2.2).

Immigrants with a non-English speaking background have lower literacy and numeracy skills

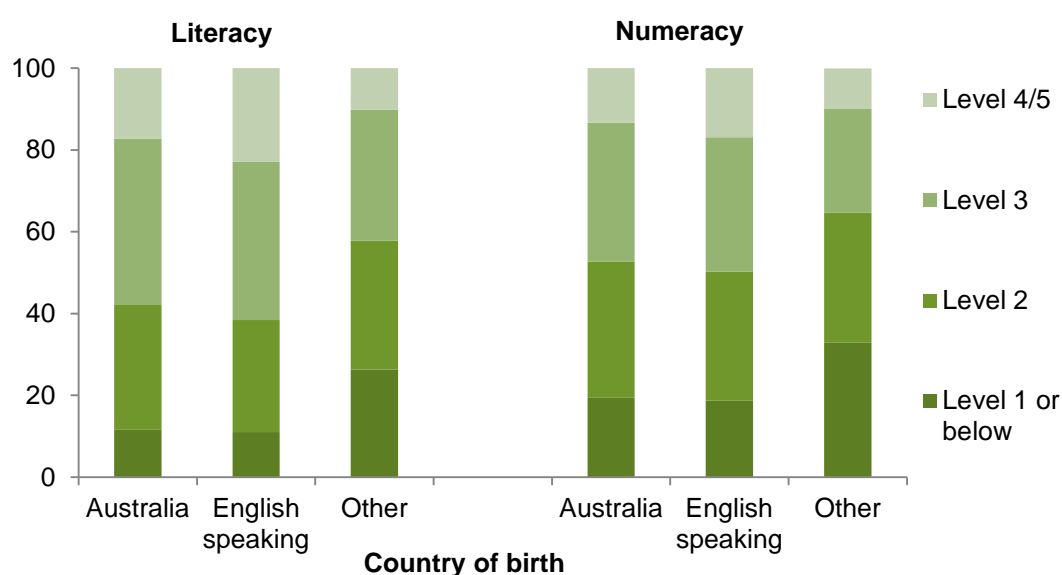
Literacy and numeracy skills are likely to be affected by one's linguistic and cultural background. Furthermore, education systems could be different across countries, producing different types of skills. Immigrants could bring skills from their country of origin that are somewhat different from those required to function in their country of residence.

Immigrants born in a 'main English speaking country' had levels of literacy and numeracy similar to those of people born in Australia (figure 2.6). Immigrants born

⁵ A comparison of educational attainment in the ALLS (2006) and PIAAC (2011–12) showed that the proportion of people with only Year 11 or lower education decreased, from about 38 per cent in 2006 to 31 per cent in 2012. This decrease allowed small increases in higher educational attainment categories (about 2 percentage points for each).

elsewhere had lower skills. For example, the proportion of immigrants born in a non-English speaking country with low (level 1 or below) English literacy was more than twice that of others.

Figure 2.6 Distribution of literacy and numeracy skill levels, by country of birth, 2011–12^a
Per cent, 15–74 year olds



^a 'Main English speaking countries' are: Canada, Republic of Ireland, New Zealand, South Africa, the United Kingdom and the United States. 'Other' are countries that are not Australia or not defined as main English speaking countries.

Source: Authors' estimates based on PIAAC data.

It is important to note that immigrants with a non-English speaking background tend to have lower *English* literacy and numeracy skills, and not low literacy and numeracy skills per se. 'The fact that immigrants, particularly those from foreign-language backgrounds, have low proficiency in the language of the assessment does not imply that they have poor proficiency in their mother tongue.' (OECD 2013a, p. 125).

Other countries that participated in the PIAAC survey reported lower literacy for immigrants compared with native-born persons. According to the OECD (2013a), the difference in mean literacy scores between foreign-born and native-born persons was relatively small in Australia, compared with most other countries. The smaller difference could be related to the emphasis placed on skills in Australian migration schemes. The OECD (2013a, p. 125) found that 'countries with points-based labour-migration schemes, such as Australia and Canada, give considerable weight to proficiency in their national languages' and that such requirements are not

possible in all countries. The greater proportion of more highly skilled migrants, the more likely differences between immigrants' skills and those of native-born people are small.

The finding that literacy and numeracy skills of immigrants born in non-English speaking countries are lower than others was also found in 2006 (Shomos, 2010).

Summary

The analysis above shows that a large share of the population has relatively low literacy and numeracy skills. People are more likely to have lower skills if they have a non-English speaking background, are more than 55 years old, or have low educational attainment. These conclusions, and the proportions of the population in each skill level, are similar to those for 2006, suggesting that differences in skills across demographic groups have not changed much.

2.3 Literacy and numeracy skills by labour market outcomes

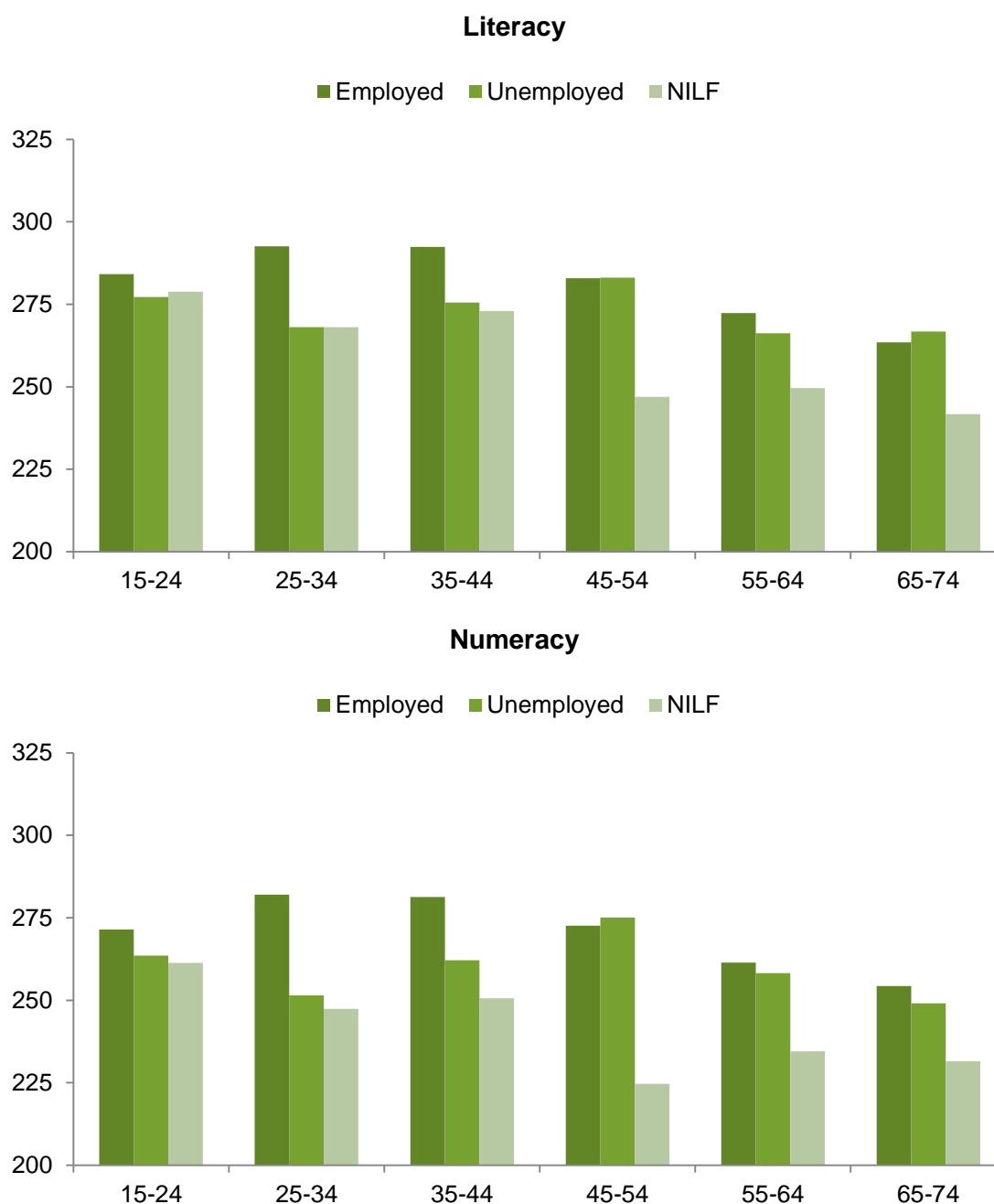
People with greater human capital (including literacy and numeracy skills) are likely to have better labour market outcomes than people with lower skills. More skilled people are more likely to participate in the workforce because their returns from working are higher than returns for people with lower levels of literacy and numeracy.

As might be expected, across all age groups, the literacy and numeracy skills of employed people are higher than the skills of people not in the labour force (figure 2.7).⁶ Shomos (2010) found a similar pattern in 2006 — people in the labour force (employed and unemployed) had higher document literacy skills than persons not in the labour force, across all age cohorts.

Differences in literacy and numeracy skills by labour force status vary across age cohorts. These differences are smallest for those under 25. Some young people who are still in education (and therefore potentially higher skilled) are not in the labour force: this contributes to raising the average scores of those not in the labour force. In contrast, for people aged 25–44 years, the differences in skills between those employed and those not employed are larger.

⁶ Note that some of these differences may not be statistically significant.

Figure 2.7 **Average literacy and numeracy test scores, by labour force status and age, 2011–12^{a, b}**



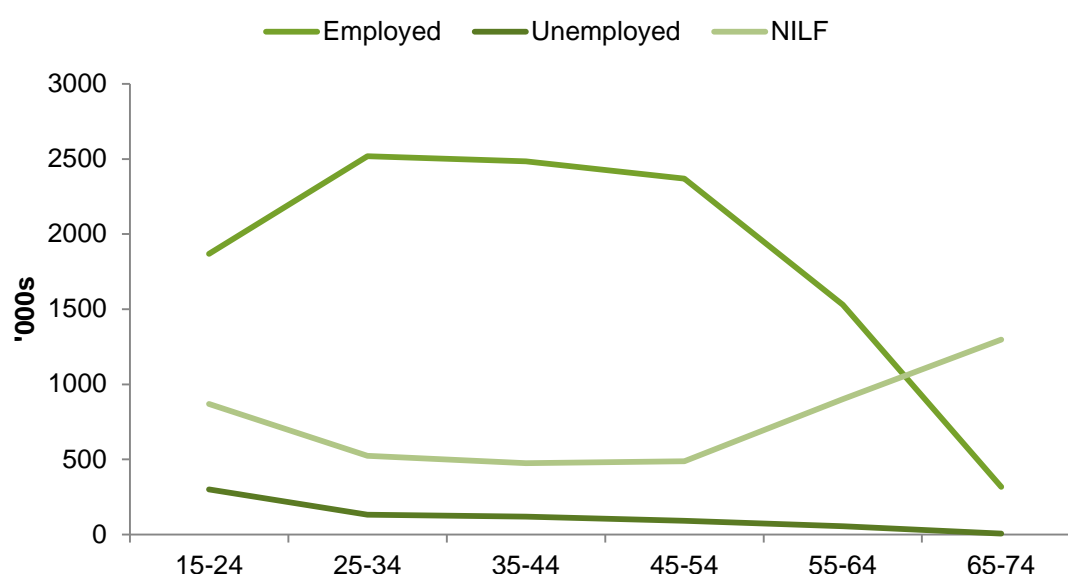
^a NILF Not in the labour force. ^b Estimates for the unemployed have large standard errors.

Source: Authors' estimates based on PIAAC data.

For employed persons, the data show a pattern of slight skill appreciation up to age 35–44, followed by a decline. This is consistent with the pattern of skill change by age depicted in figure 2.4.

For people not in the labour force a different pattern emerges. Average test scores for persons aged 15–44 years are relatively stable (about 270–275 points for literacy); average test scores are lower for persons aged 45–74 years (about 250 points for literacy). It is not clear what explains these patterns, especially why skills appear to decline so much for the 45–54 year old age group. On the other hand, it is likely that retirements of people with relatively high skills are responsible for maintaining literacy skills and increasing numeracy skills among people aged 55–74 years (figure 2.8).

Figure 2.8 Number of persons by labour force status, 2011–12^a



^a NILF Not in the labour force.

Source: Authors' estimates based on PIAAC data.

The differences in skills between the employed and those not in the labour force are typically larger for numeracy than they are for literacy. This is influenced by gender — as remarked above, women are more likely to have lower numeracy skills than men, and they are also less likely to participate.

Literacy and numeracy skills are correlated with wages. In 2011–12, the average wage of a worker with literacy skill level 1 or below was 60 to 70 per cent of the wage of a worker with skill level 4/5 (table 2.4). This is similar to the 60 per cent differential reported for 2006 by Shomos (2010).

Table 2.4 Average wage rates, by literacy skill level and gender, 2011–12
Dollars per hour,^a 15–74 year old employees

<i>Literacy skill level</i>	<i>Men</i>	<i>Women</i>
1 or below	24.64	22.55
2	27.78	24.41
3	33.01	27.93
4/5	41.64	33.85

^a Calculated using annual income from wages (not including bonuses) and hours usually worked per week in main job (see appendix B for information on how hourly wage rates were derived).

Source: Authors' estimates based on PIAAC data.

3 Modelling and results

The descriptive analysis in chapter 2 showed that literacy and numeracy skills are associated with education, age and English speaking background. In this chapter, multivariate analysis is used to isolate the links between literacy and numeracy skills and labour market outcomes (employment and wages). Section 3.1 develops a model based on the concepts set out in chapter 1 and describes the variables used in the analysis. Section 3.2 presents results and their implications.

3.1 Model and variables

Unlike descriptive analyses, which examine the correlation between two variables, econometric modelling can account for many factors to isolate the relationships between each factor and labour market outcomes.

Estimating employment and wages models

The model developed in this chapter draws on the human capital model, where labour market outcomes (such as employment or wages) are assumed to be influenced by human capital and other personal characteristics. Human capital is the embodiment of many factors that influence a person's labour market outcomes (chapter 1). But other factors can also influence employment or wages. For example, having a partner or having children may influence a person's labour force participation decision.

A human capital model of wages based on Mincer (1974) was augmented to include literacy and numeracy skills:¹

$$\ln(Wage_i) = \alpha_0 + \alpha_1 Ed_i + \alpha_2 LN_i + \alpha_3 EXP_i + \alpha_4 EXP_i^2 + \alpha_5 X_{1i} + \alpha_6 X_{2i} + \dots \alpha_{n+4} X_{ni} + \varepsilon_i$$

Where: $Wage_i$ = hourly wage rate for individual i

Ed_i = a measure of education

¹ The natural logarithm of wages is used in the model as it gives a better fit of the data.

LN_i = a measure of literacy and numeracy skills (explained below)

EXP_i = work experience

$X_1, X_2 \dots X_n$ = other factors likely to affect employment and wages

ε_i = an error term.²

The variables used are described in more detail below and in appendix B.

A similar model was estimated for employment, with labour force status (LFS) as the dependent variable.³ Three labour force states are modelled: employed, unemployed and not in the labour force.

A model which excludes literacy and numeracy skills (LN_i), will be referred to as model 1 in the remainder of the paper. The model that accounts for literacy and numeracy skills explicitly, as described above, will be referred to as model 2.

In model 2, the association between literacy and numeracy skills and wages (or LFS) is measured by α_2 .

The coefficient α_1 measures the direct association between education and labour market outcomes, once literacy and numeracy skills are taken into account. Thus, it is a measure of how important education is for developing other components of human capital that are valued in the workplace. Barrett (2012) suggested these could include other cognitive skills (critical thinking, analytic and decision-making skills) and non-cognitive skills that other researchers have noted (including perseverance and leadership skills).⁴

The relationship between educational attainment and literacy and numeracy (and other skills and knowledge), and how they jointly determine labour market outcomes can be analysed by comparing the estimates for educational attainment in each model. As Chiswick, Lee and Miller (2003) explained, if education is a good proxy for literacy and numeracy skills, then including literacy and numeracy (in

² The error term accounts for intrinsic randomness, any error of measurement or the influence of variables not explicitly included in the model.

³ For the model with labour force status, age was used instead of experience, and other factors (X_i to X_n) were slightly different (appendix C).

⁴ These cognitive and non-cognitive skills also influence educational attainment, and therefore wages and participation both directly and indirectly.

model 2) could cause collinearity⁵ problems, leading to imprecise estimates of the effects of education and literacy and numeracy skills. An alternative scenario is where there is no relationship between education and literacy and numeracy. In that case, inclusion of the literacy and numeracy variable in model 2 would not result in any major changes to the estimates in model 1. The most likely scenario, however, is that literacy and numeracy skills are developed through a range of settings, including outside of education (chapter 1). In this case, both education and literacy and numeracy skills should be included when estimating models of labour market outcomes.

Results should be interpreted as associations between literacy and numeracy skills (or education) and labour market outcomes, rather than causal effects.⁶ This and other features of the model are discussed in box 3.1.

From chapter 2, wages differ substantially between men and women. Furthermore, women with children are less likely to participate in the labour force. Therefore, models were estimated for men and women separately. The age group of interest is defined as people likely to have completed their initial post-school education and not yet retired. Therefore the sample was restricted to 25–64 year olds.⁷

⁵ Variables are collinear when they can be expressed as linear function of each other (Gujarati 1995). In practice, variables might not be highly collinear but still affect the quality of the estimates, which will be unbiased but imprecise (have high standard errors).

⁶ Causal relationships between education and literacy and numeracy skills are difficult to estimate because skills are assessed at the time of the survey, which is after education has been completed, for most people. Other researchers (Barrett 2012 using the Adult Literacy and Lifeskills Survey; Chiswick, Lee and Miller 2003 using the Survey of Aspects of Literacy) have noted the difficulties with these types of datasets in identifying instruments suitable for estimating causal relationships.

⁷ Other studies omit full-time students. The Programme for the International Assessment of Adult Competencies data does not distinguish between full-time and part-time students, so all students were retained. Alternative specifications which omitted persons who were studying and not employed did not materially affect the results.

Box 3.1 **Econometric issues with the models estimated**

Reverse causation

In the models estimated in this paper, it is assumed that no explanatory variable (including human capital) is influenced by a person's wage. In practice, wages are also likely to influence human capital investment (figure 1.1). For example, a person's income could affect how much education they undertake. If there is a positive effect of wages on human capital (education or literacy and numeracy), then the estimated coefficients for the human capital variables will be overestimates. Similarly, because a person's wage influences the decision to participate, there is the potential for reverse causality in the labour force status model, which would also lead to overestimates.

Unobserved variables

Some variables that are unobserved (such as ability) can influence wages, education or skill variables. For example, to the extent that education is positively correlated with ability, the estimated coefficients for education will be overstated. Some of the association accounts for people with higher ability undertaking more education.

Ability influences both cognitive and non-cognitive skills (chapter 1). Literacy and numeracy data may capture some of the cognitive elements of ability. Chesters, Ryan and Sinning (2013) argue that literacy and numeracy variables are likely to capture some aspects of individual ability and broader skills not observed in models that use only educational attainment.

While unobserved factors are not accounted for in this report, any bias is unlikely to change the qualitative findings:

- Leigh (2008), in a review of the literature that examined ability bias, concluded that Australian estimates for educational attainment in wages models might be biased upwards. In his analysis, Leigh concluded that ability bias meant education estimates were biased upwards by 10 per cent and adjusted results by that amount.
- In the case of labour force participation, comparing estimates that did and did not allow for unobserved heterogeneity by using panel data, Laplagne, Glover and Shomos (2007) found that the marginal effects of education for men were practically the same; for women allowing for unobserved factors reduced the effects of education by about 10 per cent.

(continued next page)

Box 3.1 (continued)**Sample selection bias**

Sample selection bias occurs when the data are not a random sample of the population. Wages are only observed for people who are employed — people who are unemployed or not in the labour force may have different characteristics. Restricting the sample to people who earn a wage could introduce biases into the estimation.

To address this, a two equations approach (so-called Heckman approach) is used in this report when estimating the wages model — a ‘selection’ (employment) equation and a ‘principal’ (wage) equation. The selection equation estimates the likelihood that a person with a given set of characteristics will be employed. The wage equation includes an adjustment factor (based on the selection equation) to estimate a wage for everybody in the sample (including those unemployed and not in the labour force). Results indicate that there is a sample selection bias for men, but not women and that the approach corrects for this (appendix C).

Variables used in the analysis

The Programme for the International Assessment of Adult Competencies (PIAAC) dataset includes many of the variables required to estimate the wage and employment models of the type described above. The variables are described briefly below, and in more detail in appendix B.

Employment and wages

The dependent variables are labour force status and the natural logarithm of wages:

- Labour force status is a categorical variable indicating whether a survey participant is employed, unemployed or not in the labour force.
- An estimate of hourly wages was derived from information on annual wages and weekly hours usually worked in a person’s main job (appendix B).

Literacy and numeracy skills

Test scores for literacy and numeracy were used as they are measured on a scale of 1 to 500, which allows for more continuous variation compared with using discrete skill levels.

There was a strong positive correlation between literacy and numeracy skills (for example, see table 2.2, presented for skill levels). To address potential collinearity issues, a variable that accounts for both literacy and numeracy skills was created

using principal component analysis (appendix B). The variable is a simple weighted average of literacy and numeracy test scores. As this variable is a measure of both literacy and numeracy, the links between either literacy *or* numeracy skills with labour market outcomes cannot be identified separately.⁸

Education

Education was measured as highest qualification attained, grouped in 4 levels:

- Year 11 or lower (this includes people with Certificates I and II)
- Year 12
- diploma or certificate (certificate is only for Certificates III and IV)
- bachelor degree or higher.

Definitions of these attainment levels are in appendix B. Year 11 or lower is the default category, meaning that estimates associated with educational attainment are measured relative to having Year 11 or lower education.

3.2 Results

Results confirm a positive association between the simple measure of literacy and numeracy skills and both employment and wages. Marginal effects⁹ for the employment and wages models, and the implications of those results, are presented in this section.

Employment

The association between literacy and numeracy and different labour market states was estimated using a multinomial logit model.¹⁰

⁸ Given the observed correlation, it would not have been possible to identify different effects in any case.

⁹ A marginal effect measures the association between a change in one explanatory variable (for example, literacy and numeracy skills or education) and the dependent variable (in this case either labour force status or wages) when all other variables are held constant.

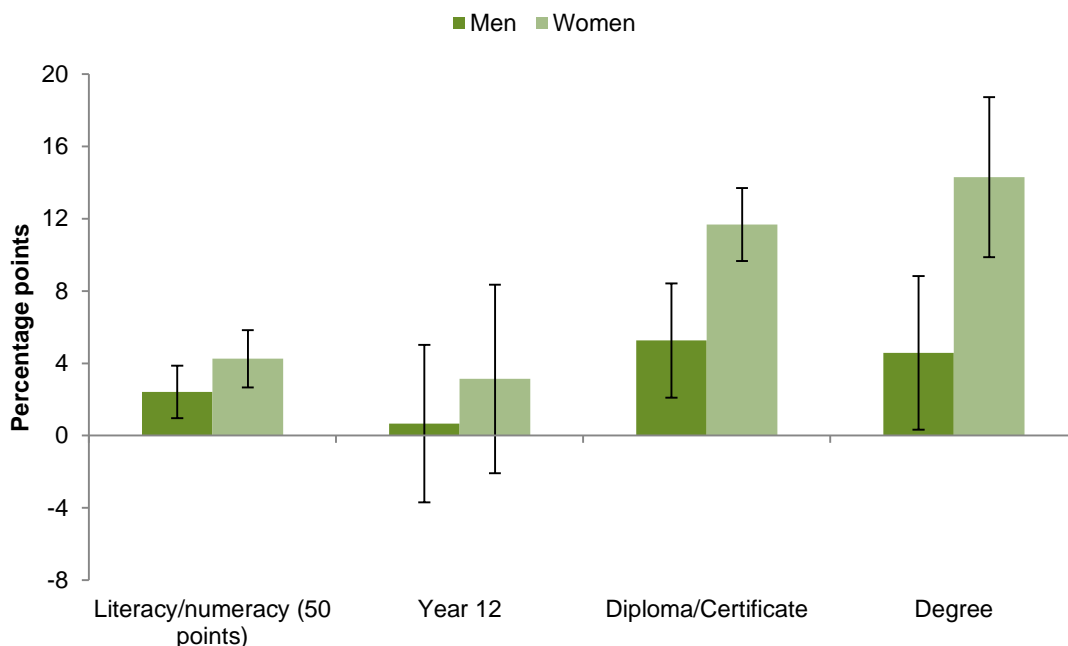
¹⁰ In the labour force status model, the dependent variable is categorical, with three mutually exclusives outcomes (employed, unemployed or not in the labour force), requiring a multinomial model.

How important are literacy and numeracy skills for employment?

The marginal effects of literacy and numeracy skills associated with the probability of employment are positive and statistically different from zero (figure 3.1). For a given level of educational attainment, and controlling for other factors, a 50-point increase in a person's average literacy and numeracy test score — equivalent to one skill level for most people¹¹ — is associated with an increased probability of employment of about 2.4 and 4.3 percentage points for men and women, respectively.

Marginal effects for educational attainment are generally positive, but much larger for women than for men (figure 3.1). This likely reflects differences in the rate of employment among men aged 25–64 years (85 per cent) and women (72 per cent). The high employment rate among men means that there is less scope to increase male employment.

Figure 3.1 **Marginal effects of literacy and numeracy and education associated with the probability of employment, 25–64 year olds^a**
Increase in probability of employment for the average person



^a Error bars indicate the 95 per cent confidence intervals.

Source: Appendix C.

¹¹ For most people, an increase in test score of 50 points is equivalent to one skill level, as the cut-off between levels are 50 points apart except for people assessed below skill level 1 (table 2.1). (Appendix A has a discussion of skill levels and test scores).

The point estimates of the marginal effects for a 50 point increase in literacy and numeracy test scores are larger than those associated with completing Year 12 (as compared to having education of Year 11 or lower).

Further tests and details on literacy and numeracy skills

The marginal effect of literacy and numeracy skills associated with the probability of employment is similar across all levels of educational attainment. That is, an increase in literacy and numeracy skills is associated with a similar increase in the probability of employment whether a person has a degree, diploma/certificate or Year 12.¹²

Results for the association between literacy and numeracy skills and employment are consistent with those reported by the Organisation for Economic Co-operation and Development (OECD, 2013a). Although that study estimated the association between literacy (not numeracy) and labour force participation, results are similar when the model is estimated on a comparable basis.¹³

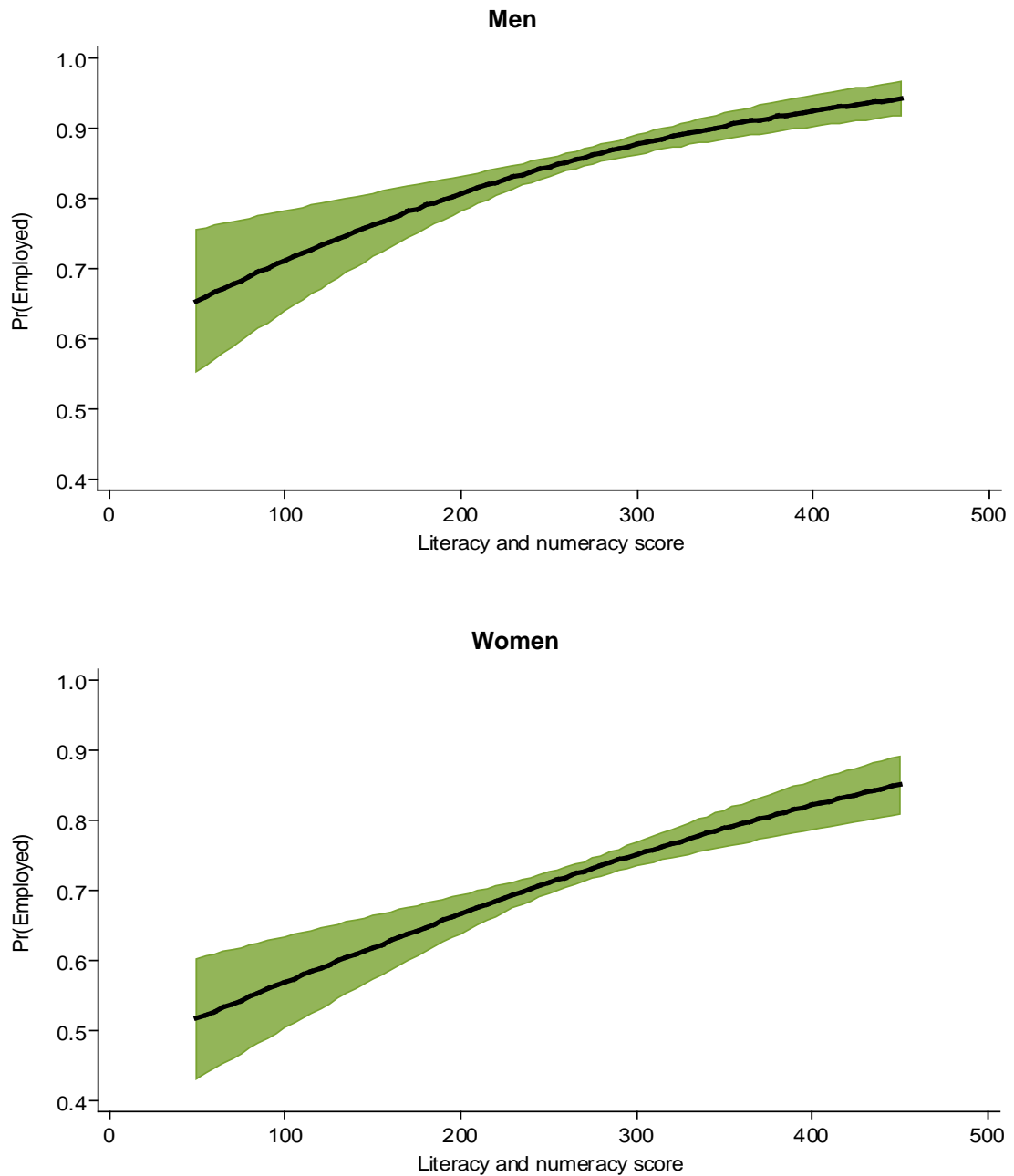
Marginal effects for literacy and numeracy skills were also estimated at different test scores to see how they change along the distribution of skills in the population. Although the relationship is depicted as non-linear, the statistical results imply that there is no difference in the *rate of increase* of the probability of being employed as skills increase (figure 3.2).¹⁴

¹² This was tested by including interaction terms for literacy and numeracy skills and educational attainment in the labour force status model. The coefficients on the interaction terms were not significantly different from zero.

¹³ The OECD (2013a) used a one standard deviation in literacy test scores and reported results using ‘odds ratios’. For comparability with the OECD (2013a), a change in literacy test scores of one standard deviation (47 points) was also used here and the model re-estimated for labour force participation. This gave an odds ratio of 1.23, very close to the 1.20 reported in OECD (2013a). An odds ratio of 1.20 means that a person is 20 per cent more likely to participate, given his or her probability of participating. That is, for someone with a 50 per cent probability of participation, an odds ratio of 1.20 implies that the person’s probability of employment increases to 60 per cent.

¹⁴ A test of significance for the marginal effects along the distribution did not reject the hypothesis that marginal effects were the same along the distribution.

Figure 3.2 Probability of employment by literacy and numeracy score^a
25–64 year olds



^a The marginal effect of a change in the probability of employment associated with a change in literacy and numeracy is estimated for different literacy and numeracy scores between 50 and 450, and is depicted by the slope of the line. Shaded areas represent 95 per cent confidence intervals.

Source: Authors' estimates using PIAAC data.

The point estimate for a change in probability of employment associated with a change in literacy and numeracy score of 50 points is larger for women (4.3 percentage points) than it is for men (2.4 percentage points). The results for

employment and unemployment can also be aggregated to the labour force participation level, so that they are comparable with Shomos (2010). When this is done, the probability of labour force participation associated with a 50 point change in test scores for women and men are 2.6 and 1.6 percentage points, respectively, similar to Shomos (2010) who reported marginal effects for women twice as large as that for men.

How do the results compare with those from models without literacy and numeracy skills?

As discussed above, comparing models both with and without literacy and numeracy skills provides insights into the relationship between educational attainment and literacy and numeracy.

The positive association between education and employment partly reflects the higher level of literacy and numeracy skills observed among those with higher educational attainment. After taking into account differences in literacy and numeracy skills, the marginal effect of improved educational attainment associated with the probability of employment decreases substantially (table 3.1). For men, the reduction in the marginal effect of a degree is over 40 per cent. For women, the decline is more modest (less than 25 per cent). That is, up to 40 per cent of the association between education and employment is attributable to literacy and numeracy skills — for example, those required to make inferences when reading texts.

Table 3.1 Marginal effects for educational attainment in the labour force status model, 2011–12^a

25–64 year olds

	<i>Men</i>		<i>Women</i>	
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>
<i>Probability of being employed</i>				
Year 12	3.2	0.7	5.5 *	3.1
Diploma or Certificate	7.3 ***	5.3 ***	14.7 ***	11.7 ***
Degree	8.3 ***	4.6 **	19.6 ***	14.3 ***
Literacy and numeracy	-	2.4 ***	-	4.3 ***
<i>Probability of being unemployed</i>				
Year 12	-0.4	-0.2	-1.0	-1.0
Diploma or Certificate	-0.3	-0.1	-1.0	-1.0
Degree	-1.6 *	-1.2	-1.1	-1.1
Literacy and numeracy	-	-0.3	-	-0.1
<i>Probability of being NILF</i>				
Year 12	-2.8	-0.5	-4.5 *	-2.2
Diploma or Certificate	-7.0 ***	-5.2 ***	-13.7 ***	-10.7 ***
Degree	-6.7 ***	-3.4 *	-18.5 ***	-13.2 ***
Literacy and numeracy	-	-2.1 ***	-	-4.2 ***

^a Marginal effects are estimates of the change in the expected probability (percentage points) of a particular labour force status associated with either a change in level of educational attainment (relative to having Year 11 or lower education), or a 50 point change in average literacy and numeracy test scores. Model 1 (without literacy and numeracy) and model 2 (with literacy and numeracy) are described above.

*** significant at 1 per cent, ** 5 per cent and * 10 per cent.

Source: Appendix C.

Other factors that might affect the association between employment and literacy and numeracy skills were also considered (results are in appendix C), including a person's English speaking background. For women, being born in a non-English speaking country increased the likelihood of being unemployed (by 2.2 percentage points). In model 1, the probability of employment among women born in non-English speaking countries is about 8 percentage points lower than for women born in Australia. After taking into account literacy and numeracy differences in model 2, this is reduced to about 5 percentage points.¹⁵ Therefore, factors other than literacy and numeracy skills also contribute to reducing the employment prospects of women from non-English speaking backgrounds.

In the case of wages, Barrett (2012) suggested there could be differences in other unmeasured productive characteristics, or possibly discrimination against immigrants of a non-English speaking background.

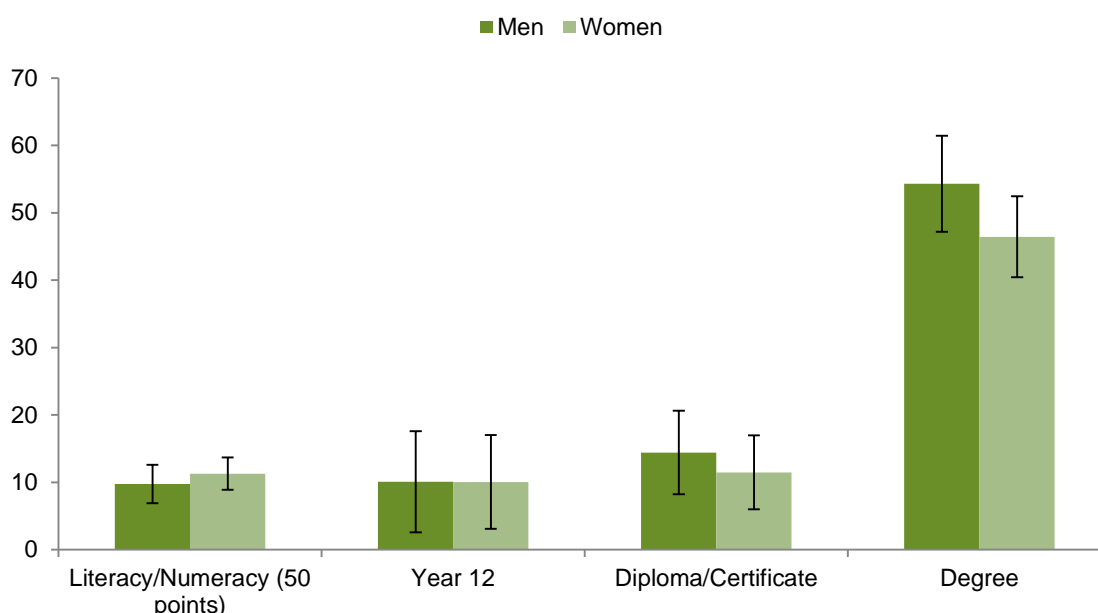
¹⁵ This marginal effect was still statistically different from zero.

Wages

How important are literacy and numeracy skills for wages?

As was the case in the participation model, an increase in adult literacy and numeracy is associated with higher wages. The marginal effects of literacy and numeracy skills are positive and statistically different from zero (figure 3.3). For a given level of educational attainment, and controlling for other factors, a 50-point increase in average literacy and numeracy score is associated with an increase in wages of about 10 per cent, for both men and women.

Figure 3.3 Marginal effects for skills and education in the wages model^a
Increase in wages, per cent



^a Error bars show the 95 per cent confidence intervals for marginal effects.

Source: Appendix C.

Marginal effects of educational attainment are positive and similar for men and women. The results imply that the positive association between raising literacy and numeracy skills by one skill level and wages is similar to that of increasing educational attainment from Year 11 to Year 12 or to diploma/certificate.

Results were similar to those of the OECD (2013a), which estimated the change in wages associated with a change in literacy¹⁶ and with results from Hanushek et al. (2013) who estimated the change in wages associated with a change in numeracy.¹⁷

How do the results compare with those from models without literacy and numeracy skills?

Estimates of the association between education and wages are reduced by about one-quarter to one-third for most levels of educational attainment, when literacy and numeracy skills are included (table 3.2). Therefore, as was the case for employment, about two-thirds to three-quarters of the association between educational attainment and wages occurs because education produces attributes other than literacy and numeracy that are valued in the workplace.

Table 3.2 Educational attainment marginal effects for wages models, 2011–12^{a, b}
25–64 year olds, per cent

<i>Explanatory variable</i>	<i>Men</i>		<i>Women</i>	
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>
Degree or higher	71.3***	54.1***	63.8***	46.6***
Diploma or Certificate	19.0***	14.5***	16.3***	11.5***
Year 12	17.3***	10.1**	14.6***	10.0***
Literacy and numeracy		9.8***		11.3***

^a Marginal effects are estimates of the increase in wages (per cent) associated with a change in the explanatory variable. Model 1 (without literacy and numeracy) and model 2 (with literacy and numeracy) are described above. ^b Estimates for educational attainment are relative to having Year 11 or lower education.

*** significant at 1 per cent, ** 5 per cent and * 10 per cent.

Source: Appendix C.

Other factors that might affect the association between employment and literacy and numeracy skills were also considered (appendix C). There is a negative association between being born in a non-English speaking country and wages. This was reduced from about -7 per cent (for men and women) to about -3 per cent for men and -2 per cent for women once literacy and numeracy skills were added to the model (appendix C). Therefore, more than half of the wage ‘penalty’ from having a

¹⁶ A version of the model was estimated with literacy only (not numeracy). A change in literacy test scores of 47 points was associated with a change in wages of 8.4 per cent, compared to 7.5 per cent reported by the OECD (2013a).

¹⁷ Hanushek et al. (2013) estimated that the change in wages associated with a one standard deviation change in numeracy skills was 10.1 per cent, for a model that pooled PIAAC results across countries.

non-English speaking background is accounted for by lower literacy and numeracy skills.¹⁸

The association between labour market experience and wages was similar in models which did and did not include literacy and numeracy skills (appendix C). Therefore, there does not appear to be a relationship between experience and literacy and numeracy in determining wages — suggesting that these skills are not acquired on the job. This is consistent with Barrett (2012) and Shomos (2010) and indicates that these empirical relationships have been stable through time.

3.3 Conclusion

Literacy and numeracy skills are an important component of human capital and higher levels of human capital are linked to better labour market outcomes. People with higher literacy and numeracy skills are, on average, more likely to participate and have higher wages than people with lower skills (chapter 2). The multivariate analysis in this chapter showed that, all else equal, an increase in literacy and numeracy of about one skill level is associated, on average, with an increased likelihood of employment of about 2 to 4 percentage points and 10 per cent higher wages.

Having estimated the potential returns associated with an improvement in skills and labour market outcomes, another question is *how* to improve literacy and numeracy.

Literacy and numeracy skills can be acquired in many ways, including during early childhood, through formal education, through on-the-job learning, and in day-to-day activities (chapter 1). This paper did not consider how important each of these factors is for developing or maintaining literacy and numeracy skills.

The results in this chapter do, however, suggest a strong link between educational attainment and literacy and numeracy skills in explaining labour market outcomes. Literacy and numeracy account for up to 40 per cent of the association between educational attainment and labour market outcomes. The results are consistent with education producing other skills and knowledge that are valued in the workplace such as higher order skills and non-cognitive skills, including perseverance and leadership (Barrett 2012).

Although education and literacy and numeracy are closely related, the analysis in chapter 2 highlighted that any increase in literacy skills since 2006 was probably

¹⁸ As noted in chapter 2, both literacy and numeracy were lower for immigrants from a non-English speaking background. This could occur because the tests were done in English.

small and numeracy skills were unchanged between 2006 and 2011–12. This was despite increases in educational attainment over that period.

People with higher levels of literacy and numeracy can understand information from dense texts and complex or abstract mathematical information. Improving literacy and numeracy is likely to lead to increases in other components of human capital, such as knowledge and higher order critical thinking. Improving these other aspects of human capital is also important for labour market success.

To summarise, literacy and numeracy skills are an important component of a person's human capital, and contribute to the development of other aspects of human capital. The modelling in this chapter demonstrated that, all else equal, there are strong links between literacy and numeracy skills and employment and wages.

A Literacy, numeracy and problem solving measures in PIAAC

The 2011–12 Programme for the International Assessment of Adult Competencies (PIAAC) is an international survey coordinated by the Organisation for Economic Co-operation and Development (OECD). It has been conducted in 23 OECD countries — as well as the Russian Federation — using survey instruments that adhere to a common set of technical and quality assurance guidelines (Caldwell and Webster 2013; OECD 2010b). The PIAAC survey complements the Programme for International Student Assessment (PISA), and follows the 2006 Adult and Literacy and Life Skills Survey, and the 1998 International Adult Literacy Survey (IALS). Much of the PIAAC survey methodology draws on these preceding surveys.

The analysis in this report is based on preliminary Australian PIAAC unit record data that were released by the Australian Bureau of Statistics (ABS) in February 2013. The survey provides information about survey participants across three separate skill domains:

- literacy
- numeracy
- problem solving in technology-rich environments.

The collection of PIAAC data in Australia is briefly described in box A.1.

An overview of the literacy, numeracy and problem solving measures in PIAAC is in section A.1. Section A.2 outlines the concept of ‘latent proficiency’ that influences the design of PIAAC. Section A.3 describes the multiple imputation approach to the subjectiveness inherent in measuring latent proficiencies and how a set of ‘plausible values’ are produced. The method for using plausible values to estimate population parameters presented in this paper is also presented.

Box A.1 **Collection of PIAAC data in Australia**

The Programme for the International Assessment of Adult Competencies (PIAAC) survey was conducted in Australia between October 2011 and March 2012, and included Australians living in private dwellings who were aged 15–74 years. People living in very remote areas or discrete Indigenous communities were not included in the survey, and children aged 15–17 years were only included with the consent of a parent or responsible adult.

The PIAAC sample was drawn from the ABS Population Survey Framework, which was built from a list of census collection districts, each of which comprises around 250 private dwellings. The collection districts used are typically the Primary Sampling Units (PSUs) used in sample selection, although in areas with low population density, PSUs are formed by grouping neighbouring collection districts. There are about 30 000 PSUs in the framework.

A multistage sample design was used: the first stage involved the selection of a sample of PSUs; the second involved the selection of ‘blocks’ within the selected PSUs; the third involved the selection of dwellings within the selected blocks; and a fourth stage involved the random selection of a person from a selected dwelling.

The sample included 11 532 households, from which 8600 respondents remained after exclusions due to scope and coverage. Of these, 8446 people completed the survey, with the remainder not completing the survey due to language or literacy difficulties.

Information was collected face-to-face, using computer-assisted interviewing. Household characteristics were collected from a responsible adult within the household and a personal interview with a randomly selected household member was used to collect demographic and other background information. The survey participant then completed the main survey instrument delivered by computer, paper or a combination of these.

Source: ABS (2013c).

A.1 Proficiency measures in PIAAC

The PIAAC survey was designed to measure population-wide proficiency levels in the domains of adult literacy, numeracy and technology-related problem solving.¹

Prior to PIAAC, the IALS from 1994–98 and the Adult Literacy and Life Skills Survey (ALLS) conducted from 2003–06 provided broad profiles of adult skills in

¹ PIAAC focuses on broad, functional concepts of literacy and numeracy. In terms of literacy, this relates to an individual’s ability to understand and use written texts to participate in society, achieve goals and develop their knowledge and potential (OECD 2012b). This is in contrast to narrower concepts of child literacy, used in assessments like the National Assessment Program for Literacy and Numeracy (NAPLAN), which focus on foundation literacy skills.

developed countries including Australia.² The PIAAC survey extends the skills framework in the previous surveys by focusing on literacy, numeracy and problem solving skills that are considered relevant to the digital age (OECD 2012b). A description of each of the three domains, as applied in PIAAC, is presented in box A.2.

Box A.2 Skill domains in PIAAC

The Programme for the International Assessment of Adult Competencies seeks to measure population proficiency levels across the domains of literacy, numeracy and problem solving in technology-related environments.

- Literacy is the ability to understand and use information from written texts in a variety of contexts to achieve goals, develop knowledge and fulfil potential. This is critical for developing higher-order skills and for positive economic and social outcomes. Previous studies have shown reading literacy to be closely linked to positive outcomes at work, social participation and lifelong learning.
- Numeracy is the ability to use, apply, interpret, and communicate mathematical information and ideas. It is an essential skill in an age when individuals encounter an increasing amount and wide range of quantitative and mathematical information in their daily lives. Numeracy is a skill *parallel* to reading literacy; it is important to assess how these competencies interact, since there are some differences in how they are distributed across the population.
- Problem solving in technology-rich environments is defined as the ability to use digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks. There is a focus on the ability to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks. In this survey, higher-order skills are identified along with basic proficiency.

Source: OECD (2012b).

Proficiency scores for each skill domain are derived on a scale ranging from 0 to 500 points. The process by which these scores are derived is briefly described in section A.2. To aid interpretation for users of the data, the scores are grouped into five skill levels for the literacy and numeracy domains, and three levels for problem solving in technology-rich environments. Level one or below indicates the lowest skill level. Skill levels represent particular sets of abilities and the thresholds are not equidistant (ABS 2013c). Tasks associated with different skill levels are outlined in tables A.1, A.2 and A.3.

² The International Adult Literacy Survey was also referred to as the Survey of Aspects Literacy in Australia (Caldwell and Webster 2013).

Table A.1 Description of skill levels in literacy PIAAC data

<i>Skill level</i>	<i>Description of task difficulty</i>
Below level 1 (0 to < 176 points)	<ul style="list-style-type: none"> • Tasks at this level require the respondent to read brief texts to locate a single piece of specific information. • There is seldom any competing information in the text and the requested information is identical in form to information in the question. • Only basic vocabulary knowledge is required, and the reader is not required to understand the structure of sentences or paragraphs. • Tasks do not make use of any features specific to digital texts.
Level 1 (176 to < 226 points)	<ul style="list-style-type: none"> • Tasks at this level require the respondent to read relatively short digital or print texts to locate a single piece of information that is synonymous with the information given in the question. • Some tasks require personal information to be entered onto a document. • Little, if any, competing information is present. • Some tasks require simple cycling through multiple pieces of information. • Knowledge and skill in recognising basic vocabulary, evaluating the meaning of sentences, and reading of paragraph text is expected.
Level 2 (226 to < 276 points)	<ul style="list-style-type: none"> • Tasks at this level require respondents to make matches between the text (which may be digital or printed) and information, and may require paraphrasing or low-level inferences. • Some competing pieces of information may be present and may require the respondent to: <ul style="list-style-type: none"> – cycle through or integrate two or more pieces of information – compare and contrast or reason about information requested in the question, or – navigate within digital texts to access-and-identify information from various parts of a document.
Level 3 (276 to < 326 points)	<ul style="list-style-type: none"> • Texts at this level are often dense or lengthy, including continuous, non-continuous, mixed, or multiple pages. • Understanding text and rhetorical structures become more central to successfully completing tasks, especially in navigation of complex digital texts. • Tasks require the respondent to identify, interpret, or evaluate one or more pieces of information, and often require inferences of different levels. • Many tasks require the construction of meaning across larger chunks of text or multi-step operations to be performed to identify and formulate responses. • Often tasks require that the respondent disregard irrelevant or inappropriate text content to answer accurately. Competing information is often present, but it is not more prominent than the correct information.

(Continued next page)

Table A.1 (continued)

<i>Skill level</i>	<i>Description of task difficulty</i>
Level 4 (326 to < 376 points)	<ul style="list-style-type: none"> • Tasks at this level often require respondents to perform multiple-step operations to integrate, interpret, or synthesize information from complex or lengthy continuous, non-continuous, mixed, or multiple type texts. • Complex inferences and application of background knowledge may be needed to perform successfully. • Many tasks require identifying and understanding one or more specific, non-central ideas in the text in order to interpret or evaluate subtle evidence-claim or persuasive discourse relationships. • Conditional information is frequently present in tasks at this level and must be taken into consideration by the respondent. • Competing information is present and as prominent as correct information.
Level 5 (376 to 500 points)	<ul style="list-style-type: none"> • At this level, tasks may require the respondent to search for and integrate information across multiple, dense texts; construct syntheses of similar and contrasting ideas or points of view; or evaluate evidence-based arguments. • Application and evaluation of logical and conceptual models of ideas may be required. • Evaluating reliability of sources and selecting key information is frequently a key requirement. • Tasks often require respondents to be aware of subtle, rhetorical cues and to make high-level inferences or use specialised background knowledge.

Source: OECD (2013a).

Table A.2 Description of numeracy skill levels in PIAAC data

<i>Skill level</i>	<i>Description of task difficulty</i>
Below level 1 (0 to < 176 points)	<ul style="list-style-type: none"> • Tasks at this level require the respondent to carry out simple processes, such as counting, sorting and basic arithmetic operations with whole numbers or money, or recognising common spatial representations in concrete, familiar contexts where the mathematical content is explicit with little or no text or distractors.
Level 1 (0 to < 226 points)	<ul style="list-style-type: none"> • Tasks at this level require the respondent to carry out basic mathematical processes in common, concrete contexts where the mathematical content is explicit with little text and minimal distractors. • Tasks usually require one-step or simple processes such as: counting; sorting; performing basic arithmetic operations; understanding simple percentages; locating and identifying elements of simple graphical or spatial representations.
Level 2 (226 to < 276 points)	<ul style="list-style-type: none"> • Tasks in this level require the respondent to identify and act upon mathematical information and ideas embedded in a range of common contexts, where the mathematical content is fairly explicit or visual with relatively few distractors. • Tasks tend to require the application of two or more steps or processes such as: calculation with whole numbers and common decimals, percentages and fractions; simple measurement and spatial representation; estimation; interpretation of relatively simple data and statistics in texts, tables and graphs.
Level 3 (276 to < 326 points)	<ul style="list-style-type: none"> • Tasks in this level require the respondent to understand mathematical information which may be less explicit, embedded in contexts that are not always familiar and represented in more complex ways. • Tasks require several steps and may involve the choice of problem solving strategies and relevant processes. • Tasks tend to require the application of number sense and spatial sense; recognising and working with mathematical relationships, patterns, and proportions expressed in verbal or numerical form; interpretation and basic analysis of data and statistics in texts, tables and graphs.
Level 4 (326 to < 376 points)	<ul style="list-style-type: none"> • Tasks in this level require the respondent to understand a broad range of mathematical information that may be complex, abstract or embedded in unfamiliar contexts. • These tasks involve undertaking multiple steps and choosing relevant problem solving strategies and processes. • Tasks tend to require analysis and more complex reasoning regarding: quantities and data; statistics and chance; spatial relationships; change, proportions and formulas. • Tasks in this level may also require comprehending arguments or communicating well-reasoned explanations for answers or choices.
Level 5 (376 to 500 points)	<ul style="list-style-type: none"> • Tasks in this level require the respondent to understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts. • Respondents may have to integrate multiple types of mathematical information where considerable translation or interpretation is required; draw inferences; develop or work with mathematical arguments or models; justify, evaluate and critically reflect upon solutions or choices.

Source: ABS (2013c).

Table A.3 Description of skill levels for problem solving in a technology-rich environment in PIAAC data

<i>Skill level</i>	<i>Description of task difficulty</i>
Below level 1 (0 to < 241 points)	<ul style="list-style-type: none"> • Tasks are based on well-defined problems involving the use of only one function within a generic interface to meet one explicit criterion without any categorical or inferential reasoning, or transforming of information. • Few steps are required and no sub-goal has to be generated.
Level 1 (241 to < 291 points)	<ul style="list-style-type: none"> • At this level, tasks typically require the use of widely available and familiar technology applications, such as email software or web browser, with little or no navigation required to access the information or commands required to solve the problem. • Problems may be solved regardless of the respondent's awareness and use of specific tools and functions. • Tasks involves few steps and a minimal number of operators. • At a cognitive level, the respondent can readily infer the goal from the task statement; problem resolution requires one to apply explicit criteria; and there are few monitoring demands (for example, the respondent does not have to check whether they have used the adequate procedure or made progress toward the solution). • Identifying contents and operators can be done through simple match; only simple forms of reasoning, such as assigning items to categories, are required. There is no need to contrast or integrate information.
Level 2 (291 to less than 341 points)	<ul style="list-style-type: none"> • At this level, tasks typically require the use of both generic and more specific technology applications. For instance, the person may have to make use of a novel online form. • Some navigation across pages and applications is required to solve the problem. • The use of tools (such as a sort function) can facilitate the resolution of the problem, and the task may involve multiple steps and operators. • In terms of cognitive processing, the problem goal may have to be defined by the person, though the criteria to be met are explicit. There are higher monitoring demands. • Some unexpected outcomes or impasses may appear. The task may require evaluating the relevance of a set of items to discard distractors. Some integration and inferential reasoning may be needed.
Level 3 (341 to 500 points)	<ul style="list-style-type: none"> • At this level, tasks typically require the use of both generic and more specific technology applications. • Some navigation across pages and applications is required to solve the problem. • The use of tools (such as a sort function) is required to make progress toward the solution, and the task may involve multiple steps and operators. • In terms of cognitive processing, the problem goal may have to be defined by the person, and the criteria to be met may or may not be explicit. There are typically high monitoring demands. • Unexpected outcomes and impasses are likely to occur. The task may require evaluating the relevance and reliability of information in order to discard distractors. Integration and inferential reasoning may be needed to a large extent.

Source: ABS (2013c).

A.2 Measuring latent proficiencies

Numeracy and literacy are estimated as ‘latent proficiencies’ — unobservable variables that describe the probability that a test respondent will respond in a particular way to a specific test item (Caldwell and Webster 2013; Foy, Galia and Li 2007). Literacy, numeracy and problem solving test items are designed to provide an indication of this latent proficiency. Proficiency scores derived from these test items are used as the basis for assigning skill levels described in section A.1.

‘Item response theory’ (IRT) is used to model the relationship between test responses and an individual’s latent proficiencies in the different skill domains (Caldwell and Webster 2013). As latent proficiency models are designed to measure underlying abilities, they are regarded as independent of the test situation from which they are generated. If this is the case, they can be used flexibly — different test items, under different situations can be administered to measure the same latent proficiency, and different questions can be posed to distinguish between individuals of different levels of proficiency.

Conceptualising skills as reflecting underlying proficiencies is in contrast with classical test models, which equate skill levels directly with test performance. Under a classical approach, a valid comparison of performance can involve only people who have undertaken the same test under the same conditions.

The measurement of latent proficiencies allows the comparison of proficiency levels among different population groups, using different questions or tasks. IRT is then used to derive a score from the response pattern to the various test questions (ABS 2013c). IRT involves the presumption that a correct response to a survey item is related to the proficiency of the respondent and the difficulty of the test item itself. The underlying proficiency can then be estimated using an approach similar to factor analysis to produce an item response function (IRF) (box A.3). The IRF forms the basis generating ‘plausible values’ for each skill domain (section A.3).

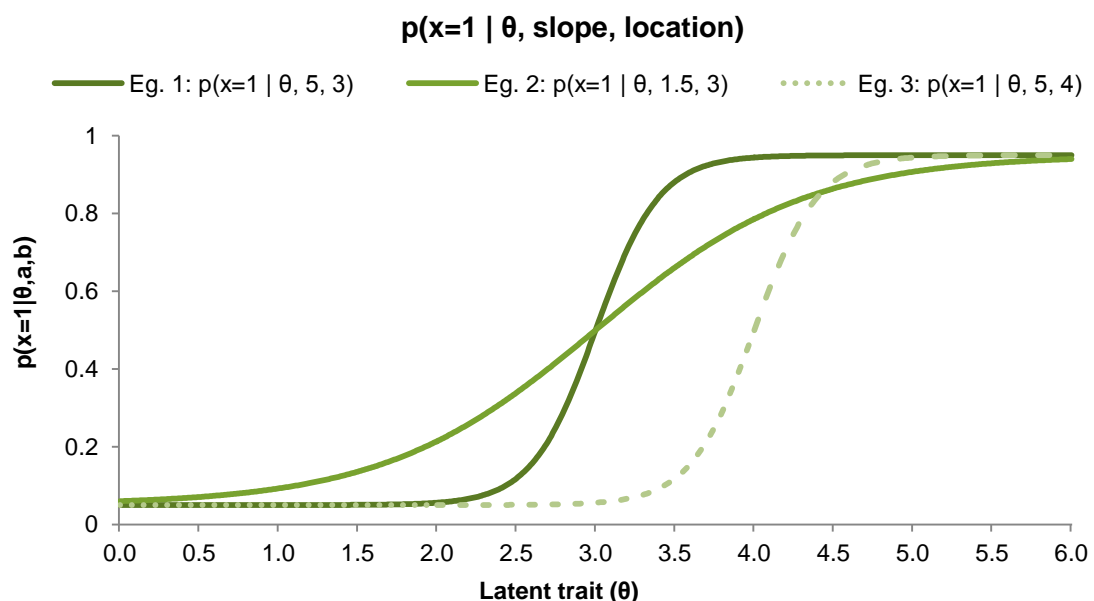
Box A.3 The basics of Item Response Theory

Item response theory (IRT) refers to a group of psychometric latent trait models that seek to measure the underlying proficiency (usually denoted by θ) that drives a test performance, rather than measuring performance in an ability test. There are three basic components of IRT.

First, the item response function (IRF) relates the latent trait to the probability of a correct response to a particular test item, given the proficiency of the individual, and the parameters of the question. The IRF is typically defined by a number of parameters drawn from a calibration process *prior* to the conduct of the survey. These parameters include:

- 'discrimination' — the steepness of the IRF. Items with a steep IRF are better at discriminating between respondents around the location (and vice-versa). In the figure below, example 1 has greater discrimination than example 2
- 'location' — the amount of the latent trait required to correctly answer an item. Different items have different locations to identify different levels of proficiency. Below, example 3 shows a location equal to 4, in contrast to examples 1 and 2, which have a location equal to 3

Variations on a four-parameter item response function for a dichotomous item^a



^a Functions are variations on the 4-parameter item response function: $p(x_j = 1 | \theta, a_j, b_j, c_j, d_j) = c_j + (d_j - c_j) / \{1 + \exp[-a_j(\theta - b_j)]\}$, where x_j is the response to item j , and equals 1 if correct and 0 if not; θ represents the latent ability; a_j the steepness of the IRF; b_j the location; c_j the lower asymptote and d_j the upper asymptote. All functions shown have a lower asymptote of 0.05, and an upper asymptote of 0.95.

(Continued next page)

Box A.3 (continued)

- 'lower asymptote' — which indicates the probability of respondents with a low level of ability 'guessing' a correct answer
- 'upper asymptote' — which indicates the probability of an incorrect response by those with a high level of ability.

The IRFs are additive, so that items can be combined to create an 'overall test response function' that relates the number of items completed, and their difficulty, to the latent trait.

IRT models assume that proficiency scores do not depend on which test items are administered. The invariance of latent proficiencies allows analysts to link different items that measure the same trait, and compare results across respondents, even if they respond to different items in the test. This allows for considerable flexibility in the design of survey instruments. This flexibility means that different survey instruments can be used to measure the same underlying abilities in different countries, cultures and languages.

Sources: Ainsworth (Nd); Mislevy, Johnson and Muraki (1992).

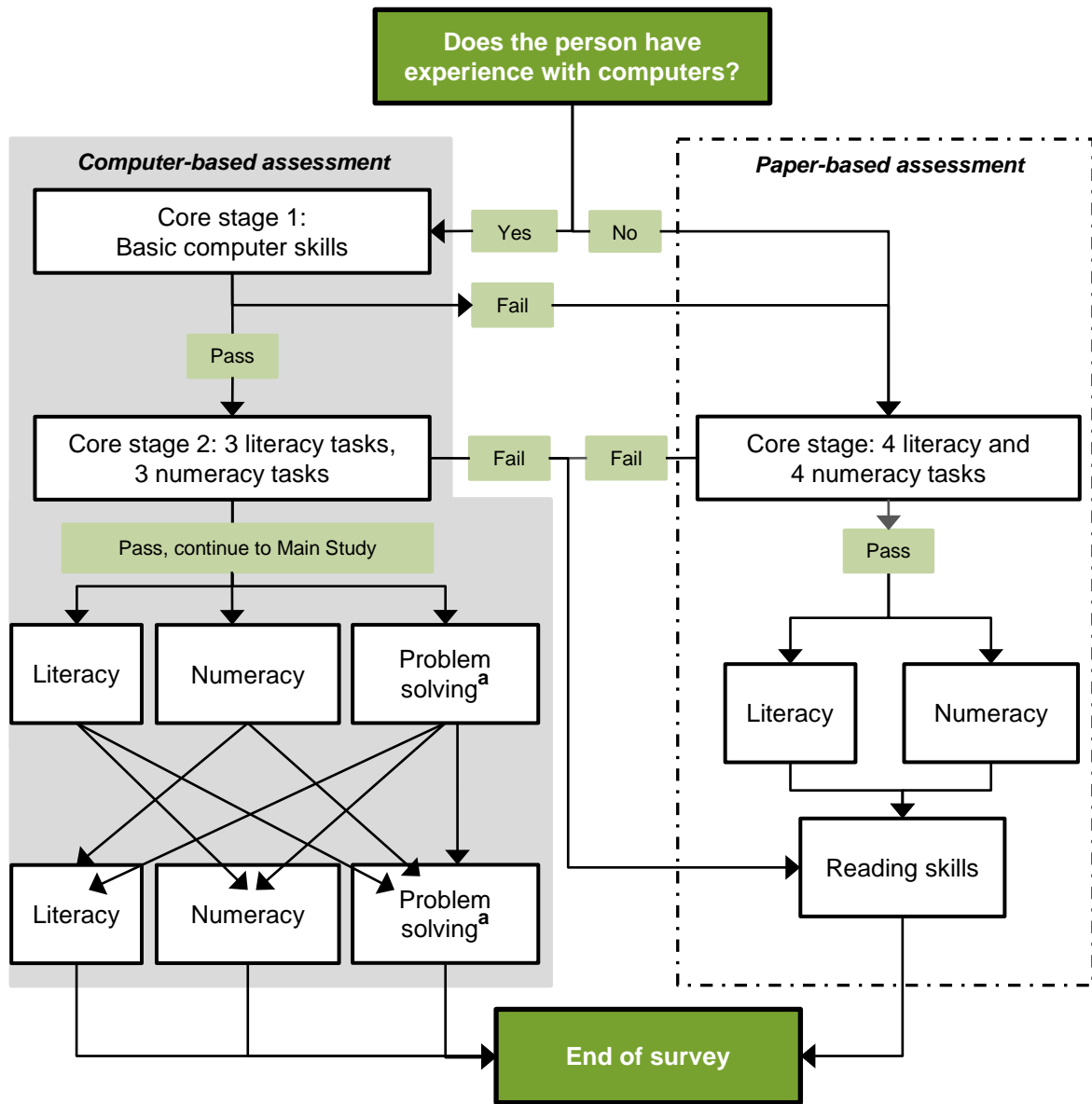
Survey design

The survey exploits the IRT approach of measuring proficiencies on the basis of a collection of test items, rather than a whole test. Respondents may be assigned different paths through the survey, with different respondents facing tasks to assess proficiency in different skill domains.

Each respondent was initially given a background questionnaire to obtain general information on education and training, employment, income and their use of skills in the domains of literacy, numeracy, and information and communication technologies. Following the background questionnaire, respondents undertook a series of *core tasks* to assess literacy, numeracy and computer skills. The tasks varied in difficulty and were based around activities that adults typically undertake in daily life, such as following written instructions, interpreting graphs, measuring with a ruler, searching on the internet or navigating websites.³ Depending on their responses to these core tasks, respondents faced different pathways through the *main tasks* to assess skills in literacy, numeracy and problem solving in a technology-rich environment (figure A.1).

³ Examples of tasks involved in the literacy components of the self-enumerated exercise can be found on the OECD website: <http://www.oecd.org/site/piaac/Literacy%20Sample%20Items.pdf>.

Figure A.1 Pathways through the PIAAC survey



^a Problem solving in a technology-rich environment.

Sources: ABS (2013c), OECD (2010).

Respondents without experience with computers or who failed a first set of ‘core tasks’ on basic computer skills were given paper-based literacy and numeracy tasks. The paper-based group were given a second set of core tasks, which consisted of a ten-minute assessment of literacy and numeracy skills. Those who did not pass the literacy and numeracy core tasks were given a test measuring basic reading skills. Those at or above a minimum standard were then randomly assigned to a 30 minute literacy or numeracy assessment, followed by a 20 minute reading assessment.

Respondents who passed both core stages were then randomly assigned to a computer-based assessment path through the *main study*. About 70 per cent of all respondents followed one of these paths. Each path involved the completion of two modules, and was expected to take around 50 minutes in total. Of those who completed the computer-based assessment:

- 50 per cent received both literacy and numeracy tasks
- 33 per cent received problem solving tasks, combined with *either* a literacy or a numeracy module
- 17 per cent received *only* problem solving tasks.

The literacy and numeracy modules were each comprised of two stages (table A.4):

- Stage 1 involved a respondent being assigned one of three testlets, each comprising two ‘blocks’ of tasks. Each respondent completed a total of nine out of a possible 18 tasks for this stage.
- Stage 2 involved a respondent being assigned to one of four testlets, each comprising either two or three ‘blocks’ of tasks. Each respondent completed a total of eleven out of a possible 31 tasks.

Depending on their responses to tasks in the preceding core stages, respondents were allocated to different ‘testlets’ that contained tasks of different levels of difficulty. The average level of task difficulty also differed between testlets. The two problem solving modules each contained 7 tasks, with each respondent required to complete all 14 problem solving tasks.

Table A.4 Literacy and numeracy module structure

Stage 1: 18 possible unique tasks — 9 tasks per testlet, each respondent takes one testlet							
Block	A1	B1	C1	D1			
Testlet 1-1	4 tasks	5 tasks					
Testlet 1-2		5 tasks	4 tasks				
Testlet 1-3			4 tasks	5 tasks			
Stage 2: 31 possible unique tasks — 11 tasks per testlet, each respondent takes one testlet							
Block	A2	B2	C2	D2	E2	F2	G2
Testlet 2-1	6 tasks	5 tasks					
Testlet 2-2		5 tasks	3 tasks	3 tasks			
Testlet 3-3				3 tasks	3 tasks	5 tasks	
Testlet 4-4						5 tasks	6 tasks

Source: OECD (2010).

A.3 Multiple imputation and plausible values

As survey respondents did not complete tasks in all three skill domains, proficiency scores are not available for all domains and individuals. To address this, ‘multiple imputation’ was used.⁴ Multiple imputation is designed to resolve the problem of missing values in datasets; it involves randomly drawing several ‘plausible’ values from a probability distribution for the missing data.

The use of a ‘split questionnaire’ that does not provide all of the desired information about each individual in the survey is a deliberate choice by survey designers. This choice is made for two reasons.

First, multiple imputation is consistent with the probabilistic nature of IRT-based testing. Even if data were not missing, multiple imputation should be used to give an indication of the spread of individuals’ proficiency scores, given their test response patterns.

Second, the use of split questionnaires can reduce the burden placed on survey respondents (ABS 2013c). Nationally representative surveys on topics such as literacy and numeracy require complex survey instruments to derive the information required — a process that involves a considerable burden on the respondent. This has the potential to affect response rates, thereby decreasing the precision of estimates. Multiple imputation approaches are a possible way of collecting information from split questionnaires, while limiting the burden on respondents (Thomas et al. 2006).

Plausible values

As proficiency scores for literacy and numeracy are estimated rather than measured, they are by definition imprecise. The lack of precision means that the scores are more appropriately viewed as being drawn from a probability distribution rather than as point estimates. The probability of people being assigned to a certain skill level relies on the questions that they are asked and their background characteristics.

An individual’s proficiency distribution can be estimated subject to a range of background characteristics; it is then referred to as the ‘posterior distribution’.

⁴ Proficiency scores for problem solving in a technology-rich environment were not imputed for respondents who undertook the paper-based survey, as they either had no computer experience, did not agree to do the test on a computer, or did not pass the computer-based core stage 1 (ABS 2013c).

The use of a distribution to represent latent proficiencies makes analysing the relationship between these proficiencies and other variables of interest difficult. The probabilistic nature of the latent proficiency measure is incorporated into the analysis by using plausible values, randomly drawn from the posterior distribution, as explanatory variables in the regression analyses presented in this paper. (This is also true of other studies using the PIAAC data).

PIAAC data include ten plausible values for literacy and ten plausible values for numeracy. Taken together, the plausible values represent a distribution of proficiency scores that reflects the uncertainty inherent in the proficiency measures.

Taken individually, single plausible values for two respondents with identical background characteristics and responses to a test are likely to differ. The extent to which they differ is affected by the extent of the uncertainty in the latent proficiency measure. As the number of imputations increases, the mean of the plausible values for two respondents with identical background characteristics and test responses converge. For this reason, plausible values cannot be used to estimate an individual's proficiency. They are suitable only for estimating proficiency at the population level.

Wu (2005) pointed out that the population distribution of proficiency scores will be smoother when using plausible values, as opposed to single point estimates.⁵ For example, if students have the same test score and the same posterior distribution, they will have slightly different plausible values. The distribution of proficiencies will be continuous, and provide a better representation of the underlying continuous distribution of proficiency in the population, than one point estimate (which is the same for both students) would.

Number of imputations

There is considerable discussion about the number of imputations required, but there is little in terms of formal recommendations. Rubin (1987) concluded that five imputations should be sufficient to obtain a valid inference, whereas Kenward and Carpenter (2007) cautioned that some analyses that involve large amounts of missing data may require 50 or more imputations. In order to reduce sampling variability, and because of increases in computational capacity, both StataCorp (2009) and Sterne et al. (2009) recommended using at least 20 imputations. PIAAC

⁵ Wu (2005) conducted a series of simulations comparing mean and variance estimates using plausible values and other methods of estimation, showing that the plausible values estimates are both unbiased and efficient.

data provides 10 plausible values. This is considered sufficient, given the random nature of the missing values and the level of data missing.

Analysis with PIAAC plausible values

As the literacy and numeracy skills data in PIAAC consist of plausible values, it is important to account for this when analysing the data. This involves two broad steps:

- Estimation of the desired parameter (for example, a population estimate or regression coefficient) is performed separately using each plausible value. This means that regression analysis using the plausible values in PIAAC requires performing ten separate regressions.
- The results of the estimation using the ten plausible values are then pooled into a single result, using Rubin's rules for multiple imputation inference (Rubin 1987, 1996).

The use of plausible values to estimate population parameters requires that variability attributable to the *imputation* of values, as well as *sampling* variability are taken into account when estimating parameters.⁶

Regression analyses using plausible values as dependent variables and background variables as independent variables produces unbiased coefficients if the posterior distribution accounts for those background variables (Mislevy, Johnson and Muraki 1992; Wu 2005). If the regressors are not included in the model that produced the plausible values, then the regression coefficients are likely to understate the true association. Importantly, the PIAAC notes produced by the ABS do not specify which background variables were used to produce the plausible values.

⁶ Variance attributable to the process of imputation is indicated by the variance of the plausible values:

$$\text{var}_{\text{imp}}(\hat{\theta}_{\text{mean}}) = \left(1 + \frac{1}{K}\right) \sum_{k=1}^K \frac{(\hat{\theta}_i - \hat{\theta}_{\text{mean}})^2}{K-1}$$

where $\hat{\theta}_i$ represents a plausible value for a latent skill, $\hat{\theta}_{\text{mean}}$ is the mean estimate of the plausible values $\hat{\theta}_i$, k represents each imputed dataset, and K is the total number of imputed datasets.

Total variance is then calculated as the sum of imputation and sampling variance (ABS 2007).

B Descriptive statistics

This appendix outlines the derivation of, and provides descriptive statistics for, variables used in the modelling of labour market outcomes in chapter 3. All data were from the Programme for International Assessment of Adult Competencies (PIAAC).

B.1 Construction of variables used in the modelling

Wages

For the wages model, an average hourly wage rate was derived from the number of hours usually worked in a person's main job and his or her annual wage income.

Responses to hours worked per week were capped at 60 hours. For some people, the number of hours worked could be more than this (and their actual wage rate would be lower than that estimated). This is unlikely to materially affect results as only 7 per cent of individuals reported working 60 or more hours per week.¹

For a large number of observations, the average hourly wage rate appeared to be too low (below the minimum wage). About 10 per cent of observations had an hourly wage below \$15. This affected relatively more men than women. Estimation results were similar when the model was re-estimated without these observations. For men, the marginal effect estimates were reduced by a small amount. The impact was negligible for women.

Education

A person's highest level of educational attainment was used in all models. Four levels of educational attainment were defined: degree or higher; diploma/certificate; Year 12; and Year 11 or lower. These were aggregated from more detailed survey responses for a person's highest level of educational attainment (table B.1).

¹ The top coding of hours worked does not appear to have substantially affected results. Limiting the sample for the wages model to include only those employed persons working less than 60 hours gave similar results to those obtained from the full sample.

Table B.1 Educational attainment variables used in the modelling

<i>Survey data response</i>	<i>Aggregated educational level of attainment</i>
Postgraduate degree, Graduate Diploma/Graduate Certificate	Degree or higher
Bachelor Degree	Degree or higher
Advanced Diploma/Diploma	Diploma/certificate
Certificate III/IV	Diploma/certificate
Year 12	Year 12
Year 11	Year 11 or lower
Year 10	Year 11 or lower
Certificate I/II	Year 11 or lower
Year 9	Year 11 or lower
Year 8 or below including never attended school	Year 11 or lower

Source: Based on the PIAAC survey data list.

One limitation of the data is that different pathways into obtaining a vocational education and training level qualification are not considered. For example, people with year 12 and with certificate III cannot be distinguished from people with less than year 12 and a certificate III.

Literacy and numeracy

Literacy and numeracy skill domains were measured in two ways in PIAAC — as test scores with a range of 0 to 500, or categorical skill levels ranging from 1 to 5. Test scores were used in the econometric modelling, to exploit the additional variation available from this measure relative to the levels measure.

The high correlation between literacy and numeracy test scores (0.90) made the use of both variables in regression-style analyses susceptible to problems of collinearity. Principal component analysis was used to combine the two skill domains into a single variable. The method uses orthogonal transformation to convert observations of (potentially) correlated variables into a set of values that are linearly uncorrelated. The first principal component ‘explained’ 95 per cent of all variation in the literacy and numeracy test scores, and placed equal weights on literacy and numeracy (table B.2). As each skill domain was equally weighted under the first component, a simple average of the literacy and numeracy test scores was used to combine the literacy and numeracy scores. In the multivariate models, this test score was divided by 50, so that marginal effects are more interpretable (marginal effects are for a change in test score of 50 points, or one skill level for most people).

Table B.2 Principal component analysis for literacy and numeracy

<i>Skill domain</i>	<i>Component 1</i>	<i>Component 2</i>	<i>Proportion explained</i>	<i>Cumulative</i>
Literacy	0.71	0.71	0.95	0.95
Numeracy	0.71	-0.71	0.05	1.00

Source: Authors' estimates based on PIAAC data.

Other variables

Other variables in the models included:

- a continuous variable indicating work experience (measured in years)
- binary variables indicating whether a person had children and their ages (0–4; 5–14 or 15–24 years)
- a binary variable indicating whether a person was living with his or her partner
- self-assessed health, on a scale from 1 (lowest) to 5 (highest)
- a person's country of birth, categorised according to whether a person was born in:
 - Australia
 - a main English-speaking country other than Australia (Canada, Republic of Ireland, New Zealand, South Africa, the United Kingdom and the United States)
 - another country.

B.2 Descriptive statistics

Table B.3 contains a description of the variables, along with their mean and standard deviation (for continuous variables).

Table B.3 Variable definition and descriptive statistics^a
25–64 year old persons

Variable	Description	Men		Women	
Labour market					
Hourly wage	(Annual wage or salary/52)/hours worked in main job	35.13	(26.73)	30.93	(31.84)
ln(wages)	Natural logarithm of hourly wage rate	3.39	(0.63)	3.28	(0.55)
Employed	Proportion of people employed	0.854		0.723	
Unemployed	Proportion of people unemployed	0.032		0.026	
Not in labour force	Proportion of people not in labour force	0.114		0.251	
Human capital					
Experience	Number of years worked	24.84	(12.00)	20.11	(11.23)
Experience squared	Number of years worked squared/100	7.61	(6.22)	5.30	(5.13)
Degree or higher	Proportion with Bachelor degree or higher	0.256		0.312	
Diploma or certificate	Proportion with Diploma or Certificate III/IV	0.375		0.307	
Year 12	Proportion with Year 12	0.135		0.116	
Year 11 or below	Proportion with Certificate I/II or Year 11 or lower	0.235		0.266	
Literacy and numeracy	Average literacy and numeracy test score (0 to 500)	278.40	(58.65)	271.50	(55.76)
Health					
	Self-assessed health status				
1	Proportion with 'poor' health	0.044		0.047	
2	Proportion with 'fair' health	0.119		0.117	
3	Proportion with 'good' health	0.334		0.290	
4	Proportion with 'very good' health	0.339		0.361	
5	Proportion with 'excellent' health	0.164		0.184	
Demographic					
Age	Age (years)	44.90	(11.27)	44.70	(11.36)
Age squared	Age squared/100	21.43	(10.14)	21.27	(10.26)
Partner	Proportion living with spouse or partner	0.618		0.604	
Child aged 0–4	Proportion with child aged 0–4	0.171		0.177	
Child aged 5–14	Proportion with child aged 5–14	0.194		0.211	
Child aged 15–24	Proportion with child aged 15–24	0.162		0.147	
Country of birth					
Australia	Proportion born in Australia	0.716		0.717	
English speaking	Proportion born in main English speaking country (not Australia)	0.118		0.113	
Other	Proportion born in non-English speaking country	0.166		0.170	

^a Standard deviations for continuous variables are shown in brackets.

Source: Authors' estimates based on PIAAC data.

C Estimation results

This appendix contains estimation output from models 1 and 2 specified in chapter 3. Models were estimated using Stata version 13.0. As explained in appendix A, all models that used literacy and numeracy data were estimated using multiple imputation commands (that is, all ten plausible values were used with parameters and their standard errors calculated according to ‘Rubin’s rules’ — see appendix A for discussion). The models were estimated using unweighted data (described in appendix B).

C.1 Labour force status models

The labour force status models were estimated using a multinomial logit model¹ with three different labour force states: employed, unemployed and not in the labour force.

The coefficient estimates in a multinomial logit model are of limited use in isolation. Therefore, only the marginal effects for women and men (aged 25–64 years) are reported in tables C.1 and C.2, respectively.² An average of marginal effects was calculated across all observations (rather than estimated at the mean of the variable).

¹ A multinomial probit model was also estimated, for comparative purposes. That model produced very similar results to the multinomial logit model.

² Coefficient estimates are available on request.

Table C.1 **Labour force status marginal effects^a**

Women aged 25–64 years

Explanatory variable	Model 1: Without literacy and numeracy						Model 2: With literacy and numeracy					
	Pr(Employed)		Pr(Unemployed)		Pr(NILF)		Pr(Employed)		Pr(Unemployed)		Pr(NILF)	
	ME	(SE)	ME	(SE)	ME	(SE)	ME	(SE)	ME	(SE)	ME	(SE)
Age	3.6***	(0.6)	-0.2	(0.3)	-3.4***	(0.6)	3.4***	(0.6)	-0.2	(0.3)	-3.2***	(0.6)
Age squared ^b	-4.9***	(0.7)	0.1	(0.3)	4.8***	(0.6)	-4.6***	(0.7)	0.1	(0.3)	4.5***	(0.6)
Education												
Year 12	5.5**	(2.7)	-1.0	(1.1)	-4.5*	(2.6)	3.1	(2.7)	-1.0	(1.0)	-2.2	(2.5)
Diploma/certificate	14.7***	(2.0)	-1.0	(0.9)	-13.7***	(1.9)	11.7***	(2.0)	-1.0	(0.9)	-10.7***	(1.9)
Degree or higher	19.6***	(2.0)	-1.1	(0.9)	-18.5***	(1.9)	14.3***	(2.3)	-1.1	(1.0)	-13.2***	(2.1)
Literacy and numeracy ^c	-		-				4.3***	(0.8)	-0.1	(0.3)	-4.2***	(0.8)
Partner	4.0***	(1.5)	-1.2**	(0.6)	-2.8*	(1.4)	3.3**	(1.5)	-1.2**	(0.6)	-2.1	(1.4)
Child aged 0-4	-32.0***	(2.3)	-0.5	(0.9)	32.5***	(2.2)	-31.7***	(2.3)	-0.5	(0.9)	32.2***	(2.2)
Child aged 5-14	-11.1***	(2.3)	1.9**	(0.8)	9.2***	(2.3)	-10.8***	(2.3)	1.8**	(0.8)	9.0***	(2.3)
Child aged 15-24	2.0	(2.4)	2.1**	(0.9)	-4.1*	(2.3)	2.6	(2.3)	2.1**	(0.9)	-4.7**	(2.3)
COB not Australia, ESB	-3.4	(2.2)	2.3***	(0.8)	1.0	(2.1)	-4.0*	(2.2)	2.3***	(0.8)	1.7	(2.1)
COB not Australia, NESB	-7.8***	(1.9)	2.2***	(0.7)	5.6***	(1.8)	-4.9**	(1.9)	2.2***	(0.7)	2.7	(1.9)
Self-assessed health status												
2 (fair)	26.0***	(4.3)	1.0	(1.7)	-27.0***	(4.2)	24.4***	(4.3)	0.9	(1.8)	-25.3***	(4.2)
3 (good)	37.4***	(3.9)	0.1	(1.5)	-37.5***	(3.9)	35.1***	(4.0)	0.0	(1.6)	-35.1***	(3.9)
4 (very good)	44.2***	(3.8)	-1.2	(1.5)	-43.0***	(3.8)	41.5***	(3.9)	-1.3	(1.6)	-40.2***	(3.9)
5 (excellent)	43.9***	(4.0)	-0.3	(1.6)	-43.6***	(4.0)	41.4***	(4.1)	-0.3	(1.7)	-41.0***	(4.0)
	n=3369	Log likelihood = -1877.61		Pseudo R ² = 0.1758			n=3369	Log likelihood = -1862.86		Pseudo R ² = 0.1823		

^a Estimates are obtained using a multinomial logistic regression model. ME denotes marginal effect, and SE is the corresponding standard error. ^b The quadratic age term has been divided by 100 to produce marginal effects of a meaningful magnitude. ^c Literacy and numeracy indicates a 50-point change in test scores.

*** significant at 1 per cent, ** 5 per cent and * 10 per cent.

Source: Authors' estimates based on PIAAC data.

Table C.2 **Labour force status marginal effects^a**

Men aged 25–64 years

Explanatory variable	Without literacy and numeracy						With literacy and numeracy					
	Pr(Employed)		Pr(Unemployed)		Pr(NILF)		Pr(Employed)		Pr(Unemployed)		Pr(NILF)	
	ME	(SE)	ME	(SE)	ME	(SE)	ME	(SE)	ME	(SE)	ME	(SE)
Age	1.3***	(0.5)	-0.1	(0.3)	-1.2***	(0.5)	1.4***	(0.5)	-0.1	(0.3)	-1.3***	(0.5)
Age squared ^b	-2.0***	(0.6)	0.1	(0.3)	1.9***	(0.5)	-2.0***	(0.6)	0.1	(0.3)	1.9***	(0.5)
Education												
Year 12	3.2	(2.2)	-0.4	(1.2)	-2.8	(1.9)	0.7	(2.2)	-0.2	(1.2)	-0.5	(2.0)
Diploma/certificate	7.3***	(1.6)	-0.3	(0.9)	-7.0***	(1.4)	5.3***	(1.6)	-0.1	(0.9)	-5.2***	(1.4)
Degree or higher	8.3***	(1.8)	-1.6*	(1.0)	-6.7***	(1.6)	4.6**	(2.2)	-1.2	(1.1)	-3.4*	(1.9)
Literacy and numeracy ^c	-		-		-		2.4***	(0.7)	-0.3	(0.4)	-2.1***	(0.6)
Partner	9.6***	(1.2)	-2.8***	(0.7)	-6.7***	(1.1)	9.2***	(1.2)	-2.8***	(0.7)	-6.4***	(1.1)
Child aged 0-4	-2.6	(2.2)	0.6	(1.1)	2.0	(2.0)	-2.8	(2.2)	0.6	(1.1)	2.2	(2.0)
Child aged 5-14	-1.1	(1.8)	2.1**	(0.9)	-1.0	(1.7)	-1.3	(1.8)	2.1**	(0.9)	-0.8	(1.7)
Child aged 15-24	1.3	(1.7)	0.5	(1.0)	-1.9	(1.5)	1.4	(1.7)	0.5	(1.0)	-1.9	(1.4)
COB not Australia, ESB	2.7	(2.0)	0.3	(1.1)	-3.0*	(1.8)	2.6	(2.0)	0.3	(1.1)	-2.9	(1.8)
COB not Australia, NESB	-3.2**	(1.6)	2.4***	(0.8)	0.8	(1.4)	-1.3	(1.7)	2.3***	(0.8)	-1.0	(1.5)
Self-assessed health status												
2 (fair)	20.7***	(4.7)	2.3	(1.7)	-23.0***	(4.5)	19.3***	(4.7)	2.3	(1.7)	-21.6***	(4.5)
3 (good)	34.0***	(4.4)	1.9	(1.5)	-35.9***	(4.3)	31.8***	(4.4)	1.9	(1.5)	-33.7***	(4.3)
4 (very good)	37.5***	(4.4)	-0.3	(1.4)	-37.2***	(4.3)	35.2***	(4.4)	-0.2	(1.4)	-34.9***	(4.3)
5 (excellent)	37.6***	(4.5)	0.5	(1.5)	-38.0***	(4.4)	35.2***	(4.5)	0.5	(1.5)	-35.7***	(4.4)
n=2928 Log likelihood= -1147.63 Pseudo R ² =0.2020 n=2928 Log likelihood =-1141.66 Pseudo R ² =0. 2061												

^a Estimates are obtained using a multinomial logistic regression model. ME denotes marginal effect, and SE is the corresponding standard error. ^b The quadratic age term has been divided by 100 to produce marginal effects of a meaningful magnitude. ^c Literacy and numeracy indicates a 50-point change in test scores.

*** significant at 1 per cent, ** 5 per cent and * 10 per cent.

Source: Authors' estimates based on PIAAC data.

C.2 Wages models

Wages models were estimated using a Heckman selection model. A Heckman model is used to control for any sample selection bias that could affect the results. Sample selection bias can occur if the data sample (in this case, employees) are not a random sample of the population. In the case of estimating wages, it is likely that people who are not employed have different characteristics to those working.

To address this, a Heckman selection model involves two steps. First, a selection equation is estimated. In this case, the selection equation is for employment. Results are used to predict the employment probability for each individual. A transformation of these predicted individual probabilities is used as an additional explanatory variable in the second stage. In the second stage, a wages equation is estimated.

To estimate a Heckman selection model, suitable ‘instruments’ are needed — variables that influence employment, but do not influence wages. The instruments used to identify the employment equation were having a child (aged 0–4; 5–14; 15–24 years) and age (including a squared term).

Results for the wages models are presented for men and women separately in table C.3. There is no selection effect detected for women, but there is a negative correlation between unobserved determinants of propensity to work and unobserved determinants of wage offers for men. This implies that, for men, unobserved factors are negatively related to working, but positively related to wages.

Table C.3 Wages model estimation output
25–64 year olds

Variable	Men				Women			
	Model 1		Model 2		Model 1		Model 2	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Selection equation								
Age	0.09***	0.03	0.11***	0.03	0.13***	0.02	0.12***	0.02
Age squared/100	-0.14***	0.03	-0.15***	0.03	-0.18***	0.03	-0.17***	0.03
Education								
Degree	0.52***	0.10	0.28**	0.13	0.75***	0.07	0.56***	0.08
Diploma/Certificate	0.36***	0.09	0.29***	0.09	0.55***	0.07	0.46***	0.07
Year 12	0.14	0.11	0.01	0.12	0.20**	0.09	0.13	0.09
Literacy/Numeracy			0.15***	0.05			0.16***	0.03
Partner	0.46***	0.07	0.46***	0.08	0.07	0.06	0.05	0.06
Child aged 0-4	-0.14	0.12	-0.17	0.12	-1.18***	0.09	-1.18***	0.09
Child aged 5-14	-0.08	0.10	-0.12	0.11	-0.39***	0.09	-0.38***	0.09
Child aged 15-24	0.10	0.10	0.10	0.10	0.15*	0.09	0.17**	0.09
COB not Australia, ESB	0.16	0.11	0.14	0.11	-0.14*	0.08	-0.16*	0.09
COB not Australia, NESB	-0.19**	0.10	-0.09	0.10	-0.32***	0.07	-0.21***	0.07
Self-assessed health								
2 (fair)	0.83***	0.16	0.82***	0.16	0.78***	0.14	0.74***	0.14
3 (good)	1.44***	0.15	1.41***	0.15	1.18***	0.13	1.11***	0.13
4 (very good)	1.62***	0.11	1.59***	0.16	1.39***	0.13	1.32***	0.13
5 (excellent)	1.53***	0.17	1.50***	0.17	1.35***	0.14	1.28***	0.14
Constant	-2.10***	0.63	-3.03***	0.70	-2.83***	0.51	-3.44***	0.53
Wage equation								
Experience	0.02***	0.00	0.02***	0.00	0.01***	0.00	0.01***	0.00
Experience squared	-0.03***	0.01	-0.03***	0.01	-0.02***	0.01	-0.01*	0.01
Education								
Degree	0.54***	0.03	0.43***	0.04	0.49***	0.03	0.38***	0.03
Diploma/Certificate	0.17***	0.03	0.14***	0.03	0.15***	0.03	0.11***	0.03
Year 12	0.16***	0.04	0.10**	0.04	0.14***	0.04	0.09***	0.04
Literacy/Numeracy			0.09***	0.01			0.11***	0.01
Partner	0.04	0.02	0.04	0.03	0.01	0.02	0.00	0.02
COB not Australia, ESB	0.07**	0.03	0.07**	0.03	-0.02	0.03	-0.03	0.03
COB not Australia, NESB	-0.07**	0.03	-0.02	0.03	-0.07***	0.03	-0.03	0.03
Self-assessed health								
2 (fair)	0.03	0.10	0.11	0.12	0.11	0.08	0.09	0.07
3 (good)	0.01	0.10	0.12	0.14	0.11	0.07	0.09	0.07
4 (very good)	0.04	0.10	0.15	0.15	0.13*	0.07	0.10	0.07
5 (excellent)	0.09	0.10	0.20	0.15	0.18**	0.08	0.16**	0.07
Constant	2.85***	0.12	2.25***	0.23	2.74***	0.10	2.24***	0.11
Lambda	-0.22***	0.05	-0.09	0.13	0.05	0.04	0.07*	0.04
Rho	-0.47***	0.11	-0.21	0.29	0.13	0.10	0.18*	0.10

*** significant at 1 per cent, ** 5 per cent and * 10 per cent.

Source: Authors' estimates based on PIAAC data.

Marginal effects in the wages models

The dependent variable in the wages model is the natural logarithm of hourly wages. Thus, the marginal effects relate the effect for a one unit change in an explanatory variable to a change in the logarithm of wages. To obtain more meaningful results, these marginal effects were converted into a percentage growth rate in wages from a change in the explanatory variable. Thornton and Inness (1989) showed that this can be estimated as follows:

$$\text{Percentage change in Wages} = X(e^{\beta}) - 1$$

Where β is the estimated marginal effect (coefficient), and X is the unit change in the dependent variable. For binary variables, such as education, $X = 1$. A one unit change for literacy and numeracy skills is implemented as an increase in test score of 50 points. These marginal effects for variables of interest are reported in table C.4.

Table C.4 Wages model marginal effects^a
25–64 year olds

<i>Variable</i>	<i>Men</i>		<i>Women</i>	
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>
Year 12	17.3***	10.1**	14.6***	10.0***
Diploma or certificate	19.0***	14.5***	16.3***	11.5***
Degree or higher	71.3***	54.1***	63.8***	46.6***
Literacy and numeracy test score		9.8***		11.3***
Experience	2.5***	2.4***	1.3***	1.0***
Experience squared	-3.4***	-3.2***	-2.0***	-1.3*
COB (other than Australia or main English speaking)	-6.5**	-2.0	-7.2***	-2.5

^a Models 1 and 2 are specified in chapter 3.

*** significant at 1 per cent, ** 5 per cent and * 10 per cent.

Source: Authors' estimates based on PIAAC data.

References

- ABS (Australian Bureau of Statistics) 2007, *Adult literacy and life skills survey: User guide 2006*, ABS Catalogue No. 4228.0.55.002, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4228.0.55.0022006?OpenDocument> (accessed 31 July 2013).
- 2013a Programme for the International Assessment of Adult Competencies, Australia, 2011-2012, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4228.02011-12?OpenDocument> (accessed 13 November 2013).
- 2013b Programme for the International Assessment of Adult Competencies, Australia, 2011-12: Explanatory notes, <http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4228.0Explanatory+Notes12011-12> (accessed 28 March 2014).
- 2013c, Programme for the International Assessment of Adult Competencies, Australia, 2011-2012: Survey methodology, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/4228.0.30.001Main%20Features2011-2012?opendocument&tabname=Summary&prodno=4228.0.30.001&issue=2011-2012&num=&view=> (accessed 25 July 2013).
- ACARA (Australian Curriculum Assessment and Reporting Authority) 2013, *NAPLAN*, <http://www.nap.edu.au/naplan/naplan.html> (accessed 8 October 2013).
- Ainsworth, A. nd, 'Introduction to Item Response Theory', www.csun.edu/~ata20315/psy427/Topic08_IntroIRT.ppt (accessed 2 August 2013).
- AWPA (Australian Workforce and Productivity Agency) 2013, 'Human capital and productivity literature review', Commonwealth of Australia, March.
- Banks, G. 2012, *Productivity policies: the 'to do' list*, presented at the Economic and Social Outlook Conference, 'Securing the Future', 1 November 2012, Melbourne.
- Barrett, G. 2012, 'The Return to Cognitive Skills in the Australian Labour Market', *Economic Record*, vol. 88, no. 280, March, pp. 1–17.
- Becker, G. 1993, 'Human Capital: a theoretical and empirical analysis', with special reference to education', 3rd ed., University of Chicago Press.
- Birch, E. 2005, 'Studies of the Labour Supply of Australian Women: What Have We Learned', *Economic Record*, vol. 81, no. 252, pp. 65–84.

-
- Cai, L. 2010, 'The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model', *Labour Economics*, vol. 17, no. 1, pp. 77–90.
- and Kalb, G. 2006 'Health status and labour force participation: evidence from Australia', *Health Economics*, vol. 15, no. 3, pp. 241–61.
- Caldwell, B. and Webster, A. 2013, 'Adult literacy surveys in Australia', *Australian Economic Review*, vol. 46, no. 1, pp. 103–109.
- Chiswick, B., Lee, Y. and Miller, P. 2003, 'Schooling, literacy, numeracy and labour market success', *Economic Record*, vol. 79, no. 245, pp. 165–81.
- Chesters, J., Ryan, C. and Sinning, M. 2013, 'The returns to literacy skills in Australia', National Vocational Education and Training Research Program Research Report.
- Cobb-Clark, D. and Schurer, S. 2011, 'The Stability of Big-Five Personality Traits', Melbourne Institute Working Paper, no. 21/11, August.
- Cutler, D. and Lleras-Muney, A. 2006, *Education and Health: Evaluating theories and evidence*, NBER Working Paper no. 12352, June.
- Department of Education, Employment and Workplace Relations (DEEWR) 2013, <http://deewr.gov.au/closing-gap-expansion-intensive-literacy-and-numeracy-programs> (accessed 1 September 2013).
- Department of Industry 2013a, *Language, Literacy and Numeracy Practitioner Scholarships Program*, <http://www.innovation.gov.au/Skills/LiteracyAndNumeracy/LanguageLiteracyAndNumeracyPractitionerScholarshipsProgram/Pages/default.aspx> (accessed 6 November 2013).
- 2013b, *Skills for Education and Employment (SEE)*, <http://www.innovation.gov.au/SKILLS/LITERACYANDNUMERACY/SEE/Pages/default.aspx> (accessed 28 May 2014).
- Desjardins, R. and Warnke, A. 2012, 'Ageing and Skills: A Review and Analysis of Skill Gain and Skill Loss Over the Lifespan and Over Time', OECD Education Working Papers, No. 72, OECD Publishing.
- Forbes, Barker and Turner 2010, *The Effects of Education and Health on Wages and Productivity*, Productivity Commission Staff Working Paper, Melbourne, March.
- Foy, P., Galia, J. and Li, I. 2007, 'Scaling the data from the TIMSS 2007 mathematics and science assessments', *TIMSS*, pp. 225–279.

-
- Green, D. and Riddell, W. 2007, 'Literacy and the Labour Market: The Generation of Literacy and Its Impact on Earnings for Native Born Canadians', cat. no. 89-552-MIE, no. 18, Statistics Canada.
- 2012, 'Understanding Educational Impacts: The Role of Literacy and Numeracy Skills' http://www.iza.org/conference_files/TAM2012/riddell_w5670.pdf (accessed 27 March 2014).
- Grossman, M. 2000, 'The Human Capital Model', published in Culyer, A. and Newhouse, J. (eds) *Handbook of Health Economics*, 2000.
- Gujarati, D. 1995, 'Basic Econometrics', 3rd edition, McGraw-Hill.
- Hanushek, E., Schwerdt, G., Wiederhold, S. and Woessman, L. 2013, *Returns to skills around the world: Evidence from PIAAC*, NBER Working Paper no. 19762, December.
- Heckman, J. and Kautz, T. 2013, *Fostering and Measuring Skills: Interventions That Improve Character and Cognition*, NBER Working paper no. 19656, November.
- Kenward, M. and Carpenter, J. 2007, 'Multiple imputation: current perspectives', *Statistical Methods in Medical Research*, vol. 16, no. 3, pp. 199–218.
- Laplagne, P., Glover, M. and Shomos, A. 2007, *Effects of Health and Education on Labour Force Participation*, Staff Working Paper, Melbourne, May.
- Leigh, A. 2008, 'Returns to Education in Australia', *Economic Papers*, vol. 27, no. 3, pp. 233–49.
- McLachlan, R., Gilfillan, G. and Gordon, J. 2013, *Deep and Persistent Disadvantage in Australia*, rev., Productivity Commission Staff Working Paper, Canberra.
- Medibank 2011 'Sick at Work, The cost of presenteeism to your business and the economy', July.
- Mincer, J. 1974, *Schooling, Experience, and Earnings*, National Bureau of Economic Research, New York.
- Mislevy, R.J., Johnson, E.G. and Muraki, E. 1992, 'Chapter 3: Scaling Procedures in NAEP', *Journal of Educational and Behavioral Statistics*, vol. 17, no. 2, pp. 131–154.
- OECD (Organisation for Economic Co-operation and Development) 2001, *The Well-being of Nations: The Role of Human and Social Capital*, Paris.
- 2010a, *Improving Health and Social Cohesion through Education*, OECD Publishing.
- 2010b, *PIAAC Technical Standards and Guidelines*, OECD Publishing.

-
- 2012a, *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*, OECD publishing.
 - 2012b, *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*, OECD Publishing.
 - 2013a, *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*, OECD Publishing.
 - 2013b, *Annex B: OECD Skills Outlook additional tables*, <http://www.oecd.org/site/piaac/annexboecdskillsoutlookadditionaltables.htm> (accessed 24 March 2014).
 - 2013c, *The Survey of Adult Skills: Reader's Companion*, OECD publishing.
- Rubin, D.B. 1987, *Multiple imputation for nonresponse in surveys*.
- 1996, 'Multiple imputation after 18+ years', *Journal of the American Statistical Association*, vol. 91, no. 434, pp. 473–489.
- SCOTese 2012, National Foundation Skills Strategy for Adults, September.
- Shomos, A. 2010, *Links Between Literacy and Numeracy Skills and Labour Market Outcomes*, Productivity Commission Staff Working Paper, Melbourne, August.
- StataCorp 2009, *Stata Multiple-Imputation Reference Manual Release 11*, StataCorp LP, College Station, Texas.
- Statistics Canada and OECD 2005, *Learning a living: First results of the Adult Literacy and Life Skills Survey*, OECD, Paris.
- Sterne, J.A., White, I.R., Carlin, J.B., Spratt, M., Royston, P., Kenward, M.G., Wood, A.M. and Carpenter, J.R. 2009, 'Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls', *BMJ: British Medical Journal*, vol. 338, <http://www.ncbi.nlm.nih.gov/pmc/articles/pmc2714692/> (accessed 23 July 2013).
- Thomas, N., Raghunathan, T.E., Schenker, N., Katzoff, M.J. and Johnson, C.L. 2006, 'An evaluation of matrix sampling methods using data from the National Health and Nutrition Examination Survey', *Survey Methodology*, vol. 32, no. 2, p. 217.
- Thornton, R. and Innes, J. 1989, 'Interpreting semilogarithmic regression coefficients in labor research', *Journal of Labor Research*, vol. 10, no. 4, pp. 443–7.
- Treasury 2010, *Intergenerational Report*, http://archive.treasury.gov.au/igr/igr2010/report/html/03_Chapter_2_Growing_the_economy.asp, (accessed 8 October 2013).

-
- Wooden, M. 2012, 'The Development of Cognitive Ability Measures in the HILDA Survey', HILDA Project Technical Paper Series, no. 1/12, June.
- Wu, M. 2005, 'The role of plausible values in large-scale surveys', *Studies in Educational Evaluation*, vol. 31, no. 2, pp. 114–128.